



Larissa Pereira Marques Cruz

**Algoritmo centrado no humano: discussão
sobre vieses de gênero em decisões de projeto
com inteligência artificial**

Tese de Doutorado

Tese apresentada como requisito parcial para obtenção do grau de Doutor pelo Programa de Pós-graduação em Design e Sociedade, do Departamento de Artes e Design da PUC-Rio.

Orientadora: Prof^a. Manuela Quaresma

Rio de Janeiro
Junho de 2024

Larissa Pereira Marques Cruz

**Algoritmo centrado no humano: discussão
sobre vieses de gênero em decisões de projeto
com inteligência artificial**

Tese apresentada como requisito parcial para obtenção do grau de
Doutor pelo Programa de Pós-graduação em Design e Sociedade
da PUC-Rio. Aprovada pela Comissão Examinadora abaixo:

Prof^a. Manuela Quaresma

Orientadora
Departamento de Artes e Design – PUC-Rio

Prof^a. Claudia Renata MontÁlvão Bastos Rodrigues

Departamento de Artes e Design – PUC-Rio

Prof. Marcos Kalinowski

Departamento de Informática – PUC-Rio

Prof. Carlos Fernando Carlim Pinto

Deloitte Consulting

Prof. Gerson Gomes Cunha

COPPE-UFRJ

Rio de Janeiro, 17 de Junho de 2024

Todos os direitos reservados. A reprodução, total ou parcial do trabalho, é proibida sem a autorização da universidade, do autor e do orientador.

Larissa Pereira Marques Cruz

Ficha Catalográfica

Cruz, Larissa Pereira Marques

Algoritmo centrado no humano: discussão sobre vieses de gênero em decisões de projeto com inteligência artificial / Larissa Pereira Marques Cruz; orientador: Manuela Quaresma. – 2024.

382 f: il. color. ; 30 cm

Tese (doutorado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Artes e Design, 2024.

Inclui bibliografia

1. Artes e Design – Teses. 2. Vieses Algorítmicos de gênero. 3. Inteligência artificial centrada No Humano. 4. Mapa mental. 5. Projeto de algoritmos inclusivos. 6. Decisões de projeto. I. Quaresma, Manuela. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Artes e Design. III. Título.

CDD: 700

A minha vó Lita e minha vó Auta por me ensinarem a potência libertadora
do estudo. (*in memoriam*)

Agradecimentos

Agradeço aos meus pais que passaram grandes perrengues para priorizar minha formação acadêmica.

Ao meu irmão que me estimulou na hora certa para seguir me desenvolvendo e minha cunhada por ser sempre atenciosa.

Aos meus irmãos-amigos Augusto, Cesar, Clarissa, Cris, Débora, Gustavo, Herberth, Karina, Livia e demais amigos que sempre estiveram de mãos dadas comigo.

Ao meu companheiro Silvio que me incentivou nas horas certas, me ajudou a revisar o texto, participou de eternos debates sobre a tese comigo e me apoiou sempre que precisei.

A Professora Manuela Quaresma, minha orientadora, pela parceria, orientações científicas e por ser muito compreensiva comigo.

Ao DAD/PUC-Rio, funcionários e professores pelo apoio e estrutura oferecida para o desenvolvimento dessa pesquisa.

Resumo

Cruz, Larissa Pereira Marques; Quaresma, Manuela. **Algoritmo centrado no humano: discussão sobre vieses de gênero em decisões de projeto com inteligência artificial**. Rio de Janeiro, 2024. 382p. Tese de Doutorado – Departamento de Artes e Design, Pontifícia Universidade Católica do Rio de Janeiro.

Este estudo aborda a iniquidade de gênero em ambientes de desenvolvimento tecnológico, apoiando-se nos pilares de tecnologia, inteligência artificial e design. Baseando-se em conceitos para caracterizar os vieses em seus efeitos de fato, como a criação de brechas digitais ou lacunas de alguns perfis de pessoas no paralelo digital, observa-se a ocorrência de vieses algoritmos de gênero, sendo criados a partir dos dados produzidos. A discussão é centrada na hipótese de que a falta de diversidade poderia limitar a compreensão e a consideração de diferentes necessidades e perspectivas, potencializando a criação de vieses inconscientes nos algoritmos e exacerbando as disparidades digitais. Propõe-se que seja possível tangibilizar marcadores ético-sociais que os especialistas em Inteligência Artificial consideram ao desenvolver algoritmos, sendo categorizados pela temática da estruturação técnica, aspectos de UX, princípios de IA e experiências pessoais. A pesquisa se volta para o indivíduo decisor, parte do projeto, suas justificativas e modelos mentais. Inicialmente de forma individual e posteriormente de forma coletiva, mas mantendo uma identidade de grupo por gênero, nesse caso, feminino e masculino. Foram realizadas 20 entrevistas com o método de inquirição contextual em que os especialistas simulam a construção de um algoritmo para divulgação de um anúncio de emprego em tecnologia. Para aprofundar a discussão, foram realizados dois grupos de foco para avaliar se o uso de direcionadores baseados em conceitos de Design focados em inclusão, pode gerar um aumento no envolvimento e na profundidade de suas análises, promovendo uma tomada de decisão mais consciente. Em projetos de grande impacto social, essa abordagem pode contribuir para um desenvolvimento mais sensível e inclusivo e auxiliar na mitigação de vieses de gênero.

Palavras-chave

Vieses Algorítmicos de gênero; Inteligência artificial centrada No Humano; Mapa mental; Projeto de algoritmos inclusivos; Decisões de projeto.

Abstract

Cruz, Larissa Pereira Marques; Quaresma, Manuela (Advisor). **Human-centered algorithm: gender biases discussions in artificial intelligence design.** Rio de Janeiro, 2024. 382p. Tese de Doutorado – Departamento de Artes e Design, Pontifícia Universidade Católica do Rio de Janeiro.

This study addresses gender inequity in technological development environments, relying on the pillars of technology, artificial intelligence, and design. Based on concepts to characterize biases and their actual effects, such as the creation of digital gaps or voids for certain profiles in the digital parallel, the occurrence of gender algorithm biases, created from the data produced, is observed. The discussion is centered on the hypothesis that a lack of diversity could limit the understanding and consideration of different needs and perspectives, potentially enhancing the creation of unconscious biases in algorithms and exacerbating digital disparities. It is proposed that it may be possible to materialize ethical-social markers that artificial intelligence experts consider when developing algorithms, categorized by themes of technical structuring, UX aspects, AI principles, and personal experiences. The research focuses on the decision-maker, part of the project, their justifications, and mental models. Initially individually and later collectively, but maintaining a group identity by gender, in this case, female and male. Twenty interviews were conducted using the contextual inquiry method in which experts simulate constructing an algorithm for the dissemination of a job advertisement in technology. To deepen the discussion, two focus groups were held to evaluate whether the use of drivers based on Design concepts focused on inclusion could lead to increased engagement and depth of their analyses, promoting more conscious decision-making. In projects of significant social impact, this approach can contribute to more sensitive and inclusive development and help mitigate gender biases.

Keywords

Algorithmic Gender Biases; Human Centered Artificial Intelligence; mind map; Inclusive Algorithmic Design; Algorithmic Design.

Sumário

1	Introdução	15
1.1	Tema: Vieses algoritmos de gênero	16
1.2	Problema de pesquisa	17
1.2.1	Reconhecimento do problema	21
1.2.2	Delimitação do problema	22
1.3	Questões de Pesquisa	25
1.4	Hipótese	25
1.5	Objeto da pesquisa	25
1.6	Objetivos	26
1.6.1	Objetivo geral	26
1.6.2	Objetivo específico	26
1.7	Justificativa	27
1.8	Metodologia	27
1.9	Estrutura da tese	28
2	Relações digitais e a reprodução da desigualdade de gênero	31
2.1	A organização em rede e o mapeamento dos algoritmos	31
2.2	A tecnologia da informação em tempos de economia da atenção	34
2.3	Paralelo digital: consciência algorítmica e a participação democrática nas redes	37
2.4	Brechas Digitais: reprodução de desigualdades e o outro digital	39
2.5	Aspectos práticos para a desigualdade: Segregação Vertical de Gênero	41
3	A tecnologia e o design de Algoritmo centrado no humano: questões de projeto	43
3.1	Aspectos técnicos	43
3.2	Auditando Algoritmos e outras técnicas de transparência	45
3.3	Elementos e princípios da Inteligência Artificial Centrada no Humano: um framework para desenvolvimento de sistemas inteligentes	47
3.4	Questões para o Design do modelo de IA design de algoritmo centrado no humano	50
3.5	Design inclusivo, Design Universal e Design para todos	52
4	Metodologia	55
4.1	Mapeamento Sistemático	55
4.1.1	Estratégia de busca	55
4.1.2	Seleção de artigos	56
4.1.3	Extração de Dados	56
4.2	Entrevistas semiestruturada com adição do método de Inquirição Contextual e formulário online	60
4.3	Grupo de foco	68
5	Análise de Resultados	72
5.1	Experiências Pessoais	72

5.2	Estrutura Técnica	73
5.3	Princípios de Inteligência Artificial	75
5.4	Aspectos de UX	76
5.5	Guia de Diretrizes para Algoritmos Sensíveis	78
5.6	Análise das entrevistas: Experiências Pessoais	79
5.7	Análise das entrevistas: Estrutura Técnica, Aspectos de UX e Princípios de Inteligência Artificial	83
5.7.0.1	Planejamento: Definição de Requisitos	84
5.7.0.2	Planejamento: Dados utilizados	88
5.7.0.3	Desenvolvimento: Método e desafios técnicos	96
5.7.0.4	Testes, métricas e validação	103
5.7.0.5	Feedbacks e Iterações	108
5.7.1	Efetividade de resultados e Impacto na contratação	113
5.7.2	Guia de Diretrizes para Algoritmos Sensíveis	118
5.7.3	Mapa mental dos marcadores	121
6	Discussão do Guia de Diretrizes para Algoritmos Sensíveis.	124
6.1	Discussão por indicadores	124
6.1.1	1. Considere a diversidade sob a perspectiva das interações humanas e considere a maneira como as pessoas mudam ao longo do tempo	128
6.1.2	2. Identifique vieses relacionados à capacidade, gênero, raça e/ou acessos e preveja possíveis interações incorretas que possam acontecer	131
6.1.3	3.Crie uma diversidade de caminhos para a participação em sua experiência	133
6.1.4	4.Assuma que o projeto terá incertezas, necessidade de iterações e adaptações e trabalhe com a utilização de habilidades complementares	135
6.1.5	5.Desenvolva uma alfabetização básica em acessibilidade e se aprofunde nos critérios de acessibilidade específicos relevantes para suas soluções	136
6.1.6	6.Adote uma definição mais flexível de projeto. Abra os processos e convide pessoas com habilidades relevantes a contribuir com seu projeto, mesmo que não sejam programadores(as).	137
6.2	Votação dos indicadores	138
6.3	Observações finais	139
7	Conclusão	141
7.0.1	Limitações e oportunidades para futuros trabalhos	144
A	Apêndice: Termo de Consentimento Livre e Esclarecido	152
B	Apêndice: Entrevistas	157
B.1	Afonso	157
B.2	Bruno	170
B.3	Caio	181
B.4	Davi	191
B.5	Francisco	204

B.6	João	217
B.7	Pedro	224
B.8	Renato	231
B.9	Roberto	238
B.10	Vinicius	247
B.11	Alice	258
B.12	Ana	265
B.13	Denise	272
B.14	Joana	279
B.15	Leandra	283
B.16	Luciana	293
B.17	Mariana	304
B.18	Rebeca	317
B.19	Regina	325
C	Transcrição: Grupo de Foco	339
C.1	Grupo De Foco (homens)- Checklist Para Desenvolvimento De Algoritmos Sensíveis	339
C.2	Grupo de foco (mulheres) - Checklist para o desenvolvimento de algoritmos sensíveis	359

Lista de figuras

Figura 1.1	Gráfico retirado do artigo: New evidence of Facebook's sexist algorithm (Global witness, 2023)	24
Figura 1.2	Esquemático da metodologia. Fonte: figura elaborada pela autora	29
Figura 3.1	Esquemático de auditoria de códigos	46
Figura 3.2	Ecossistema de Agentes para desenvolvimento de sistemas inteligentes com técnicas de HCD	52
Figura 5.1	Categorias: estrutura de mapa mental	73
Figura 5.2	Esquema desenvolvido pela autora: relação entre Princípios de IA e Aspectos de UX.	77
Figura 5.3	Dados demográficos dos especialistas. Fonte: Figura elaborada pela autora	80
Figura 5.4	Formação acadêmica dos especialistas	81
Figura 5.5	Experiência profissional dos especialistas	82
Figura 5.6	Esquemático da análise de resultados	84
Figura 5.7	Trecho 1. Figura elaborada pela autora.	85
Figura 5.8	Trecho 2. Figura elaborada pela autora.	86
Figura 5.9	Trecho 3. Figura elaborada pela autora.	86
Figura 5.10	Trecho 4. Figura elaborada pela autora.	86
Figura 5.11	Trecho 5. Figura elaborada pela autora.	87
Figura 5.12	Trecho 6. Figura elaborada pela autora.	87
Figura 5.13	Trecho 7. Figura elaborada pela autora.	87
Figura 5.14	Trecho 8. Figura elaborada pela autora.	89
Figura 5.15	Trecho 9. Figura elaborada pela autora.	89
Figura 5.16	Trecho 10. Figura elaborada pela autora.	90
Figura 5.17	Trecho 11. Figura elaborada pela autora.	90
Figura 5.18	Trecho 12. Figura elaborada pela autora.	91
Figura 5.19	Trecho 13. Figura elaborada pela autora.	91
Figura 5.20	Trecho 14. Figura elaborada pela autora.	92
Figura 5.21	Trecho 15. Figura elaborada pela autora.	93
Figura 5.22	Trecho 16. Figura elaborada pela autora.	93
Figura 5.23	Trecho 17. Figura elaborada pela autora.	94
Figura 5.24	Trecho 18. Figura elaborada pela autora.	94
Figura 5.25	Trecho 19. Figura elaborada pela autora.	95
Figura 5.26	Trecho 20. Figura elaborada pela autora.	95
Figura 5.27	Trecho 21. Figura elaborada pela autora.	96
Figura 5.28	Trecho 22. Figura elaborada pela autora.	97
Figura 5.29	Trecho 23. Figura elaborada pela autora.	98
Figura 5.30	Trecho 24. Figura elaborada pela autora.	98
Figura 5.31	Trecho 25. Figura elaborada pela autora.	98
Figura 5.32	Trecho 26. Figura elaborada pela autora.	99
Figura 5.33	Trecho 27. Figura elaborada pela autora.	100
Figura 5.34	Trecho 28. Figura elaborada pela autora.	100

Figura 5.35 Trecho 29. Figura elaborada pela autora.	101
Figura 5.36 Trecho 30. Figura elaborada pela autora.	101
Figura 5.37 Trecho 31. Figura elaborada pela autora.	102
Figura 5.38 Trecho 32. Figura elaborada pela autora.	102
Figura 5.39 Trecho 33. Figura elaborada pela autora.	103
Figura 5.40 Trecho 34. Figura elaborada pela autora.	104
Figura 5.41 Trecho 35. Figura elaborada pela autora.	105
Figura 5.42 Trecho 36. Figura elaborada pela autora.	105
Figura 5.43 Trecho 37. Figura elaborada pela autora.	106
Figura 5.44 Trecho 38. Figura elaborada pela autora.	106
Figura 5.45 Trecho 39. Figura elaborada pela autora.	106
Figura 5.46 Trecho 40. Figura elaborada pela autora.	107
Figura 5.47 Trecho 41. Figura elaborada pela autora.	107
Figura 5.48 Trecho 42. Figura elaborada pela autora.	108
Figura 5.49 Trecho 43. Figura elaborada pela autora.	108
Figura 5.50 Trecho 44. Figura elaborada pela autora.	109
Figura 5.51 Trecho 45. Figura elaborada pela autora.	109
Figura 5.52 Trecho 46. Figura elaborada pela autora.	110
Figura 5.53 Trecho 47. Figura elaborada pela autora.	110
Figura 5.54 Trecho 48. Figura elaborada pela autora.	110
Figura 5.55 Trecho 49. Figura elaborada pela autora.	111
Figura 5.56 Trecho 50. Figura elaborada pela autora.	112
Figura 5.57 Trecho 51. Figura elaborada pela autora.	112
Figura 5.58 Trecho 52. Figura elaborada pela autora.	112
Figura 5.59 Trecho 53. Figura elaborada pela autora.	113
Figura 5.60 Trecho 54. Figura elaborada pela autora.	113
Figura 5.61 Trecho 55. Figura elaborada pela autora.	114
Figura 5.62 Trecho 56. Figura elaborada pela autora.	115
Figura 5.63 Trecho 57. Figura elaborada pela autora.	116
Figura 5.64 Trecho 58. Figura elaborada pela autora.	116
Figura 5.65 Trecho 59. Figura elaborada pela autora.	118
Figura 5.66 Trecho 60. Figura elaborada pela autora.	118
Figura 5.67 Trecho 61. Figura elaborada pela autora.	119
Figura 5.68 Trecho 62. Figura elaborada pela autora.	120
Figura 5.69 Trecho 62. Figura elaborada pela autora.	121
Figura 5.70 Gráfico comparativo entre unidades de análise registrada em todas as entrevistas do grupo de homens vs grupo de mulheres. Figura elaborada pela autora.	121
Figura 5.71 Mapa mental - Análise de resultados	122
Figura 6.1 Guia de Diretrizes para Algoritmos Sensíveis	125
Figura 6.2 Painel para preenchimento dos questionamentos debatidos no grupo de foco acerca de cada direcionador do Guia.	127
Figura 6.3 Votação dos indicadores - grupo de foco com participantes homens	139
Figura 6.4 Votação dos indicadores - grupo de foco com participantes homens	139

Lista de tabelas

Tabela 4.1	Critérios de Inclusão e exclusão	56
Tabela 4.2	Formulário de Extração de Dados	57
Tabela 4.3	Categorização	59
Tabela 5.1	Códigos para Experiências Pessoais	73
Tabela 5.2	Códigos para estrutura Técnica. Fonte: desenvolvida pela autora.	74
Tabela 5.3	Códigos para Princípios de IA. Fonte: Tabela elaborada com conteúdo retirado de Fjeld et al (2020)	76
Tabela 5.4	Códigos para Aspectos de UX. Fonte: Tabela elaborada com base em Hassenzahl(2008)	78
Tabela 5.5	Códigos do Guia de Diretrizes para Algoritmos Sensíveis	79
Tabela 6.1	Perfil dos participantes: grupo de foco com mulheres	126
Tabela 6.2	Perfil dos participantes: grupo de foco com homens	126

Lista de Abreviaturas

IA – Inteligência Artificial

TI – Tecnologia da Informação

UX – User Experience

HCI – Human Computer Interaction

HCD – Human Centered Design

NLP – Natural Language Processing

HAI - Human Centered Artificial Intelligence

1

Introdução

As organizações em rede propiciam novas maneiras de interação, aproximam e constroem pontes, mas também podem emular realidades que reproduzem comportamentos culturais mais arraigados sem considerar diversidades. Uma vez que a comunicação se torna descentralizada e generalista (Castells, 1996) os usuários das inúmeras ferramentas, plataformas, redes sociais passam a ser também fontes produtoras de conteúdo. Denominados por Smart People (Comité Invisível, 2015) as pessoas/cidadãos que também existem nos mundos digitais funcionam como “receptores e geradores de ideias, serviços e soluções”.

Desta forma, o uso generalizado da internet gera uma grande massa de dados. Para extrair informação dessa grande base é aplicada uma técnica denominada Big Data que se trata da análise e interpretação semântica de um enorme volume de dados (Mayer-Schonberger e Cukier, 2013). Esta técnica está localizada sob um grande eixo temático conhecido por Inteligência Artificial (IA), que engloba métodos mais específicos para a interpretação e modelagem de dados, como, por exemplo, Aprendizado de Máquina - ML/Machine Learning (Russel, 1995).

Não é de hoje, nem novidade da última década, que IA e ML são tecnologias conhecidas e vem sendo desenvolvidas a medida que a importância dos dados prevalece e seu acesso populariza. Segundo Russel (1995) Aprendizado de Máquina é um método de análise de dados que automatiza o desenvolvimento de modelos analíticos. Usando algoritmos que aprendem iterativamente a partir de dados, o Aprendizado de Máquina permite que os computadores encontrem insights ocultos sem serem explicitamente programados para procurar algo específico. No entanto, essa técnica semiautomática, responsável por reconhecer padrões nos dados, vem influenciando crescentemente as opiniões e comportamentos do dia a dia das pessoas que utilizam as redes.

A inteligência das máquinas é construída através do aprendizado gerado pela observação dos dados apresentados. Os algoritmos agentes do processamento inteligente têm capacidade de interpretar e aprender com a grande massa de dados, que por sua vez, podem refletir conceitos estereotipados, vieses de gênero, sociais, de classe e raça perpetuando esses comportamentos em

vieses algorítmicos (Kusner et al., 2017). Os algoritmos emergentes estão nascendo sem considerar, por exemplo, décadas de pesquisas e teorias que trazem reflexões e definições de como a linguagem pode estar apropriada de termos sexistas (Leavy, 2018). Conforme (Agarwal et al., 2018) um dos problemas fundamentais inerentes à lógica dos sistemas computacionais é a incapacidade de separar as informações que processam do contexto em que são geradas.

Algoritmos possuem o poder de direcionar afirmativamente quais informações, dados e cenários nos serão apresentados, atuando como um forte recurso para a nossa participação em contextos de vida pública online. Utilizamos massivamente os sites de busca, que nos permitem acessar um volume incrível de base de dados, e nos dão alcance mundial a diversos tipos de informação (Gillespie, 2014). Por sua vez, os algoritmos de recomendação por trás dos websites, mapeiam nossas preferências, desenhando a peculiaridade de cada perfil para poder sugerir assuntos, dados, propagandas de maneira assertiva (Ganhör et al., 2022)

Nas redes sociais, algoritmos gerenciam nossas interações, destacando conteúdos e informações de perfis de amigos ou perfis comerciais que temos afinidade em detrimento de outros que em algum momento escolhemos que estivessem em nosso círculo seletivo de conteúdo. Algoritmos de recomendação são projetados para calcular o que é “quente” e o que está “no topo” das discussões das plataformas e agrupar dados por similaridade que estão sendo produzidos sem limites (Kumar e Tanwar, 2020).

Juntos, os algoritmos com distintas finalidades nos viabilizam caminhos para participar de discursos sociais e políticos, impulsionando o sentimento de inclusão e pertencimento a essas comunidades digitais. Portanto, os algoritmos são uma chave lógica para a distribuição do sensível quando, a partir do controle do fluxo das informações, há o manejo para a percepção das informações pelos usuários (Langlois, 2013).

1.1

Tema: Vieses algoritmos de gênero

Após a modelagem do projeto de ML, o algoritmo roda pela base de dados selecionada (*dataset*) e vai extraindo dos dados as características necessárias para responder o problema colocado no início do projeto. A depender da metodologia selecionada no ato de projeto, as decisões tomadas pelo algoritmo poderão ter mais ou menos visibilidade. (Kim e Doshi-Velez, 2021), conforme aprofundado em item 3.2 do capítulo 3.

No entanto, é no encontro entre o algoritmo, suas premissas e o *dataset* selecionado que os dados geram resultados e são direcionados a responder os

problemas iniciais. Desta forma, vieses pessoais podem ser inseridos de forma inconsciente nos projetos de modelo algorítmico (Ali et al., 2019).

Vieses algorítmicos são os erros que podem ser interpretados a partir de uma perspectiva da análise crítica dos resultados advindos da classificação automática de dados (Ali et al., 2019). Estes resultados, quando comparados a um recorte da realidade e contextualização dos dados de entrada, podem acusar que o aprendizado sobre os dados está gerando algum desvio.

Os vieses podem ser intencionais ou não, podendo ser direcionados a partir da modelagem do problema, podem refletir decisões mais intrínsecas da pessoa que está tomando decisões ao montar o modelo automático ou mesmo na reflexão do aprendizado sob o *dataset* (Melchiorre et al., 2021). Os dados são o coração de qualquer aplicação de IA. Visto que o algoritmo não consegue aprender a partir daquilo que ele não “enxerga”, os resultados podem refletir *gaps* reproduzidos pelos dados, ou mesmo alguma falta no momento da montagem da coleta de dados ou na etapa de limpeza e estruturação dos dados que consiste na preparação do *dataset* (Gillespie et al., 2014).

Para cada tipo de aplicação, sendo ela baseada em visão computacional, textos, áudios, existem diferentes manifestações de vieses. Por exemplo, na pesquisa realizada por Buolamwini e Gebru (2018), observa-se que os sistemas de reconhecimento facial eram altamente imprecisos na classificação de rostos de mulheres negras. No artigo pioneiro - Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification, as autoras mostraram ainda que o modelo era mais preciso para pessoas que se identificavam como homens e de pele branca (Buolamwini e Gebru, 2018). O viés de gênero tem sido amplamente estudado em sistemas de Processamento de Linguagem Natural (PLN), pois essa é a forma mais evidente de viés de gênero (Sun et al., 2019). Em particular, nos sistemas de tradução de idiomas, modelos de IA atribuem pronomes a profissões de acordo com estereótipos de gênero; por exemplo, os modelos atribuem automaticamente os pronomes ele/dele a profissões como médicos e pilotos, enquanto atribuem ela/dela as enfermeiras e comissárias de bordo (Cho et al., 2021) e a partir dessa inferência, alguns conteúdos podem ser direcionados somente para aqueles usuários que o algoritmo classificou como público-alvo.

1.2

Problema de pesquisa

E por que tanto se discute sobre os impactos em potencial que algoritmos podem trazer aos indivíduos e para suas relações na sociedade? Algoritmos tem a capacidade de fazerem recortes de conteúdos que serão direcionados às

pessoas. Enquanto você navega por alguma rede social, você está produzindo dados via cliques e engajamentos. Esses dados podem representar características bem pessoais, como gostos específicos de consumo, preferências musicais, filosóficas e outras coisas mais. A partir do desenho do seu perfil, ou como o algoritmo está interpretando os seus dados, você será submetido a conteúdos direcionados conforme o “juízo” dos algoritmos de recomendação. Desta forma, os pesos e medidas adotados por esses recursos inteligentes interferem no alcance de informações para cada indivíduo (Bellamy et al., 2018).

Portanto, voltemos ao instante do surgimento do projeto do algoritmo em que pessoas estão por trás das decisões. Durante o projeto de modelagem, pontuamos três momentos importantes de tomadas de decisão que podem trazer imparcialidades do ponto de vista técnico (Kalinowski et al., 2023). No primeiro, há que se escolher métodos e regras específicas derivadas da estatística que possuem uma lógica para serem empregados na escrita do código. No segundo, há a contextualização do problema e a depender da temática, incorrerão questões específicas que o especialista deverá definir. E o terceiro momento de decisão se dá por possíveis inferências de lógica e de escolha do conjunto de *features* que serão utilizados.

Cada pessoa viveu uma experiência de vida diferente, em contextos de sociedade diferentes, possuem características físicas, psíquicas que tornam o ser humano único. Tiveram exposições temáticas e interesses diferentes. Difícil dizer que as decisões técnicas, mesmo que apoiadas por métodos comprovados, estariam estanques a qualquer influência sociocultural do executor do projeto. Portanto, o contexto importa?

Conforme a Pesquisa Nacional por amostra de domicílio (Pnad) do IBGE, apenas 20% do total de aproximadamente 580 mil profissionais que atuam com tecnologia no Brasil em 2022, são mulheres. Diante desse cenário, a representação majoritariamente masculina em ambientes de desenvolvimento de tecnologia pode gerar silenciosamente um movimento de reforço de uma perspectiva social mais recortada.

No amplo contexto da implementação de ML e IA, um modelo é considerado tendencioso em relação ao gênero se o desempenho e/ou resultado do modelo for parcial contra um segmento da população com base em seu gênero (Wang et al., 2022). No entanto, o viés de gênero é mais prejudicial quando não é tão facilmente perceptível. Especialmente nos sistemas de recomendações online, programas sociais, sistemas de justiça e policiamento, que implementam algoritmos de ML/AI, as decisões tomadas pelos sistemas automatizados podem ser tendenciosas em relação ao gênero, mas não há uma maneira definitiva de confirmar isso (Wang et al., 2022). Além disso, os siste-

mas de ML/AI utilizam geralmente o conceito binário de gênero, o que não reflete a realidade do mundo e ignora completamente a população que pode não se identificar com a identidade masculina ou feminina (Barbosa e Chen, 2019).

São colocados questionamentos centrais para entender a profundidade dos vieses gerados, sendo eles refletidos pelo *dataset*, mas também podendo ser reflexos dos parâmetros inseridos na modelagem do algoritmo. Para mitigar a persistência de vieses na formulação da modelagem de dados, há como incorporar mecanismos que fomentem ações combativas? Entendendo esta ação de programação como um ato projetual, meios de gerar transparência e explicabilidade podem auxiliar a aproximação do problema identificado com a solução adequada.

É necessário que os mecanismos de fomento de produção e consumo de dados por interfaces digitais ganhem luz para que todo o fluxo de interação e a experiência de uso sejam mais transparentes e inteligíveis. Os algoritmos fazem parte da experiência da existência digital de milhares de pessoas, as estratégias de reconhecimento de padrões devem considerar o desenvolvimento da consciência algorítmica dos usuários. Essa contextualização implicará diretamente no impacto informacional que os sistemas automáticos podem gerar, reflexões que compõe o conceito de “algoritmo crítico” (Beer, 2017).

Existem alguns princípios para ajudar a compreensão humana sobre os mecanismos de sistemas de IA, denominados sistemas com explicabilidade (eXplainable AI - XAI) ou sistemas XAI. Esses sistemas são aplicados para trazer entendimento em uma linguagem acessível ao ser humano sobre as capacidades dos sistemas com IA e evidenciar o que foi feito para a criação dos modelos, o que eles realizam, o que irão realizar (Gunning et al., 2019). Recentemente as questões éticas decorrentes do uso das aplicações de Inteligência Artificial, fizeram a temática da transparência ressurgir com mais força, trazendo a tona discussões sobre falta de confiabilidade e vieses gerados (Agarwal et al., 2018).

A denominação de responsabilização (*accountability*) é definida como a capacidade de verificar se um sistema de IA está atuando conforme as premissas declarada e o que é necessário para determinar se o sistema tem alguma atuação em possíveis desvios gerados, conhecidos como culpabilidade (*blame-worthiness*) (Kim e Doshi-Velez, 2021). No entanto, as regras de negócio estabelecidas para alcançar a transparência podem alterar o desempenho e robustez para diversos sistemas baseados em IA.

Accountability em IA representa um conjunto de técnicas que podem auxiliar a indicar se o sistema está sendo justo (*fairness*) em relação ao contexto de aplicação. Existem várias categorias de abordagens, sendo algumas

delas: transparência (dados, processos e software de código aberto); modelos interpretáveis; inspeção pós-hoc das saídas do modelo; desempenho empírico (pré e pós-mercado); e propriedades garantidas pelo projeto (Kim e Doshi-Velez, 2021).

No entanto, é difícil gerar uma definição universal de justiça algorítmica devido à complexidade dos múltiplos parâmetros cuja classificação depende da contextualização social e cultural (Wang et al., 2022). Diante desse tema emergente, muitas propostas de estudo vem surgindo.

A base da conceituação de justiça algorítmica é dividida em dois grupos para análise, com recorte individual e em grupo. Do ponto de vista individual o caminho de análise é justiça através da conscientização (Fairness Through Awareness) (Dwork et al., 2012), em que indivíduos com perfis similares devem ser expostos a resultados similares sob a medição em que “um algoritmo é justo quando retorna predições similares a indivíduos similares”. Já a concepção de justiça através do desconhecimento (Fairness Through Unawareness) traz a definição de que um algoritmo é justo desde que quaisquer atributos protegidos, como raça, idade, sexo, não sejam explicitamente usados no processo de tomada de decisão (Gajane e Pechenizkiy, 2017). Por fim, a terceira dimensão da perspectiva individual de análise é denominada Justiça Contrafactual, em que a decisão é justa em relação a um indivíduo se for a mesma tanto no mundo real quanto em um mundo contrafactual onde em hipótese o indivíduo pertencesse a um grupo demográfico diferente (Kusner et al., 2017).

Com recorte em grupo, a Justiça Algorítmica considera a divisão da população em grupos, definindo um grupo modelo com a seleção de atributos protegidos como raça, raça mais gênero e busca garantir que as medidas estatísticas dos resultados sejam iguais entre grupos/subgrupos. Os resultados comparativos são balizados por quatro pilares, sendo paridade demográfica, probabilidades equalizadas, igualdade de oportunidades e justiça diferencial. essas classificações levam em consideração conceitos estatísticos para garantir mais equidade de oportunidades de resultados (Foulds et al., 2020). Desta forma, a apuração dos atributos interfere diretamente na formação dos grupos protegidos e por consequência no desdobramento dos resultados. Embora a maioria das medidas de justiça algorítmica seja definida para tarefas de classificação, existem algumas medidas voltadas para sistemas de recomendação, denominadas Injustiça de Não Paridade (Foulds et al., 2020).

Em sistemas de recomendação (SRs), algoritmos de filtragem colaborativa fornecem recomendações para usuários (consumidores), baseando-se principalmente nas interações usuário-item coletadas, por exemplo, ao ouvir músicas ou assistir filmes. Embora os dados de interação não contenham explíci-

tamente informações sobre atributos protegidos do usuário, como gênero, raça ou idade, um modelo ainda pode codificar informações sensíveis em suas representações. Esses vieses codificados nos modelos podem levar ao fortalecimento de “bolhas de filtro” baseadas nas demografias dos usuários (Bellamy et al., 2018) e intensificar os vieses sociais existentes nos dados, aumentando assim a injustiça do SR (Buonocore, 2019).

Há o estabelecimento de uma conexão entre justiça e viés em algoritmos e o contexto mais amplo de justiça e viés na sociedade (Foulds et al., 2020). Os autores trazem como base o conceito de interseccionalidade baseado na definição de Kimberlé Crenshaw que desenvolveu o conceito de interseccionalidade para se referir às múltiplas forças sociais, identidades sociais e instrumentos ideológicos e a necessidade de considerar todos esses fatores de maneira conjunta, e não separada na análise de direitos sociais (Crenshaw, 2013). Emerge o conceito de justiça interseccional para IA, é proposto o método de análise abrangente e interdisciplinar sobre como determinados critérios se relacionam com as preocupações de campos variados, incluindo humanidades, direito, privacidade, economia e aprendizado de máquina estatístico (Foulds et al., 2020).

Portanto, a contextualização, o entendimento dos cenários e não somente dos atributos, é um ponto de extremo valor para a análise de toda a jornada dos algoritmos.

1.2.1

Reconhecimento do problema

A partir da premissa que o usuário é parte produtora e deveria ter mais ferramentas de domínio para o entendimento dos sistemas, fala-se sobre a corresponsabilidade na interação com processos automáticos. Desta forma, toda a possibilidade de acesso a ferramentas, apoio ao desenvolvimento de habilidades e infraestrutura, como velocidade da internet e acesso a dispositivos apropriados, deverão ser considerados. As diferenças de acessos, denominada Divisão/Exclusão Digital (Dutton Reisdorf, 2019) podem, portanto, acarretar impactos sociais e democráticos que devem ser considerados no desenho dos cenários digitais.

De modo que toda a realidade de acessos e conjunturas sociais são espelhadas em proporções diretas ao mundo digital, assim como vemos ao vivo, são criados espaços inatingíveis para aqueles que não representam o modelo ideal e que ocupam minoritariamente os lugares de formação, atuação, planejamento e produção de tecnologia. Dentre outras desigualdades, são criadas as brechas digitais de gênero (Wacjman, 2006), representando tal déficit de participação feminina nas áreas tecnológicas, culminando na formação do

outro digital (Gordano, 2009), a persona que tem suas necessidades e desejos negligenciados, prejudicando determinadas experiências digitais.

Todas essas questões refletem diretamente na formação do panorama da construção social-digital. Princípios éticos no uso e elaboração da Inteligência Artificial devem trazer elementos como transparência e garantia de segurança que refletirão nos arranjos mais democráticos e igualitários. Fazendo um paralelo com a arquitetura da informação, nessa tese serão elencados alguns componentes para a possibilidade de formação de um modelo mais acertado e menos excludente de sistemas automáticos coparticipativos.

1.2.2

Delimitação do problema

Consideremos o sistema humano – máquina (algoritmo) – humano, em que o projeto de algoritmo é estruturado por um ser humano, com suas limitações e particularidades e em decorrência disso, a máquina configurada terá interações variadas e por fim os humanos serão impactados em diferentes profundidades. É possível verificar correlação entre as decisões de projeto, a configuração da “máquina” e os efeitos e impactos gerados no ser humano que é alcançado se tornando parte do sistema? Esse rastreamento é de natureza mais abstrata e menos comprobatória. Se por um lado analisar os dados gerados e realizar experimentos comparativos foca no desdobramento da tecnologia e características do ecossistema digital específico, essa pesquisa está estruturada para identificar e analisar de maneira qualitativa a profundidade e desdobramentos das decisões parciais tomadas pelos especialistas que modelam o cerne de sistemas inteligentes, os algoritmos.

Conforme o entendimento que os algoritmos aprendem com dados de situações do passado para prever a adequação de futuros cenários (Mayer-Schonberger e Cukier, 2013), vamos fazer um recorte temático para elucidar de maneira mais concreta a situação em que um algoritmo automatiza ações que efetivamente agravam vieses de gênero.

Consideremos uma situação em que um anúncio deve ser publicado em meios digitais para atrair candidatos a uma vaga de emprego. Pesquisas mostram que métodos de contratação algorítmicos podem aumentar a precisão da previsão em mais de 50% em comparação com métodos tradicionais.

Porém, investigações comprovaram a disparidade do alcance de anúncios automáticos para a busca de candidatos para determinadas vagas de emprego (Buonocore, 2019) e o reforço de vieses de gênero em sistemas de recomendação (Dev e Phillips, 2019; Dodge et al., 2019; Dwork et al., 2012). Além disso, a falta de geração de oportunidades de emprego para determinados grupos,

reflete diretamente nas questões de equidade profissional.

Pesquisas vem sendo direcionadas para equacionar a maneira em que os anunciantes que contratam os serviços de anúncio definem seus requisitos e as características dos candidatos que o algoritmo retorna para os selecionados para o direcionamento do conteúdo. A enorme eficácia financeira das plataformas de publicidade online deve-se em parte às características de segmentação precisa dos usuários. Algumas das evidências já geradas, levam ao entendimento que alguns usuários são invisibilizados (Bucher, 2012) por não manterem uma constância participativa nas redes, e por sua vez, tornando sua individualidade passível de desaparecimento virtual.

Somado a isso, o alcance e efetividade desses planos de divulgação podem levar a uma entrega de anúncios potencialmente discriminatória, mesmo quando os anunciantes definem seus parâmetros de segmentação como altamente inclusivos. Ali et al., (2019) demonstrou que essa entrega enviesada ocorreu no *Facebook*, devido a efeitos de otimização de resultados, bem como às próprias previsões da plataforma sobre a relevância dos anúncios para diferentes grupos de usuários. O estudo comprovou que tanto o orçamento do anunciante quanto o conteúdo do anúncio contribuem significativamente para o viés na entrega de anúncios do *Facebook*. De forma crítica, foi observado um viés na distribuição de anúncios relacionados a oportunidades de emprego e habitação, envolvendo questões interseccionais como gênero e raça, para a definição de alcance dos anúncios (Ali et al., 2019).

Recentemente, houve mais uma investigação que comprovou a disparidade do alcance de anúncios de vagas de emprego e ainda com recorte específico de vagas para tecnologia. Em março e abril de 2023, A ONG Global Witness contratou o serviço do Facebook para publicação de anúncios contendo links para vagas de emprego em várias áreas e profissões no Reino Unido, Holanda, França, Índia, Irlanda e África do Sul, incluindo empregos para pilotos de companhias aéreas, entregadores, cabeleireiros, psicólogos e trabalhadores de TI (Witness, s.d.). Do resultado, o que ficou mais marcado foi o contrabalanceamento dos seguintes alcances: 91% das vagas para mecânicos foram visualizadas por homens e 79% das vagas para professores de pré-escola foram visualizadas por mulheres. De maneira mais detalhada estão as comparações dos alcances dos anúncios por recorte de vagas e gênero na França e Holanda conforme gráfico 1.1.

Essas descobertas revelam como o Facebook está exibindo anúncios de vagas de emprego de forma substancialmente direcionada por recorte de gênero. Segundo a Global Witness (Witness, s.d.) para todos os anúncios foi determinada apenas uma restrição: os anúncios deveriam ser exibidos para

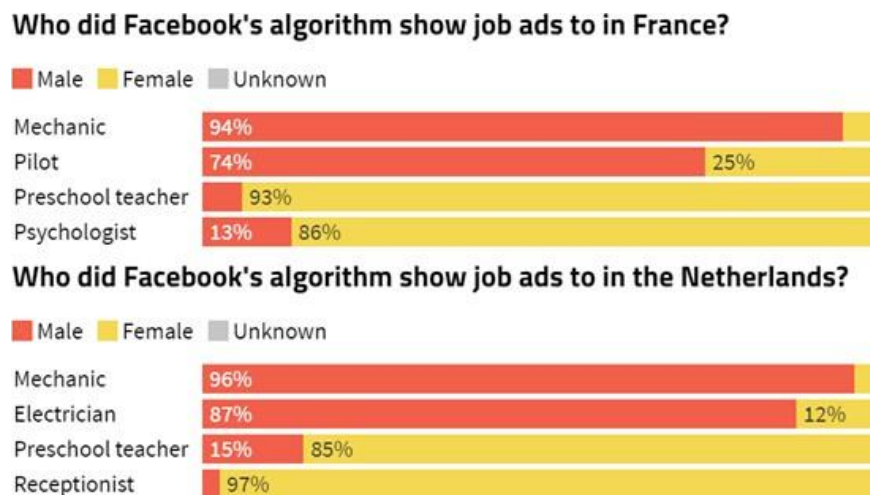


Figura 1.1: Gráfico retirado do artigo: New evidence of Facebook's sexist algorithm (Global witness, 2023)

adultos que moravam ou tinham estado recentemente em qualquer um dos países selecionados. Desta forma, a organização relata que os usuários que foram alcançados foram escolhidos conforme direcionamento do algoritmo da empresa.

Motivada pela constatação de que mecanismos como esses, em que a tecnologia reproduz preconceitos intrínsecos a estrutura de sociedade vigente, a presente investigação faz uma análise de cenário nesse contexto. Consideremos como três principais atores desse mecanismo: a empresa que contrata os serviços, o profissional que modela a solução e o alcance do produto medido em resultados de conversão. Propomos a investigação concentrada no segundo ator para entender a contextualização das decisões de projeto que profissionais podem tomar para modelar tais soluções.

Para esta pesquisa, vamos considerar, de uma maneira tácita, a observação de pontos de decisão que possivelmente estão carregados de influências do contexto do sujeito decisor.

A partir do entendimento do modelo mental das decisões de especialistas durante o desenvolvimento de algoritmos, será avaliado se há diferenças nas tomadas de decisões entre profissionais dos gêneros masculino e feminino que podem gerar ou mesmo agravar vieses algoritmos. Essa avaliação ficará no campo tácito, a medida em que termos associados a questões sociais e de gênero surjam durante a declaração das decisões tomadas, a possibilidade de conexão a alguma situação de iniquidade será analisada.

Além disso, vamos investigar se o uso de um checklist baseado em conceitos de design inclusivo pode eticamente melhorar o processo de desenvolvimento de algoritmos, considerando as diversas experiências dos usuários

1.3

Questões de Pesquisa

A contextualização das questões sócio demográficas que impactam o desenvolvimento pessoal de quem executa a modelagem do algoritmo pode acarretar direcionamentos diversos nas estratégias de projeto e desdobrar em impactos para a iniquidade de gênero. Portanto, as principais questões de pesquisa que orientam essa investigação são:

1. (QP1) Será que as diferentes experiências e vivências sociais de cada especialista em IA impactam na geração do modelo mental aplicado para o desenvolvimento de algoritmos, facilitando um cenário de replicação digital de vieses de gênero?
2. (QP2) É possível identificar os principais componentes dos modelos mentais que os(as) especialistas empregam ao projetar algoritmos, com atenção aos seus respectivos gêneros?

1.4

Hipótese

É possível identificar marcadores que revelam se especialistas em IA consideram aspectos ético-sociais em suas decisões ao desenvolver algoritmos. Introduzir direcionadores de Design focados em inclusão durante discussões com esses especialistas, pode gerar um aumento no envolvimento e na profundidade de suas análises, promovendo uma tomada de decisão mais consciente. Em projetos de grande impacto social, essa abordagem contribui para um desenvolvimento mais sensível e inclusivo. Acredita-se que a falta de diversidade em geral, as limitações na compreensão e consideração de diferentes perspectivas e necessidades pode agravar a geração de vieses inconscientes nos algoritmos e ampliação das brechas digitais.

1.5

Objeto da pesquisa

O objeto desta pesquisa compõe-se nos processos e modelos mentais que especialistas em IA utilizam para a criação de algoritmos. Considera-se especialista, nesse caso, pessoas que trabalham profissionalmente desenvolvendo algoritmos para sistemas de IA.

1.6

Objetivos

1.6.1

Objetivo geral

O objetivo geral dessa pesquisa é mapear se as diferentes perspectivas de gênero (masculino e feminino) geram impactos no projeto de algoritmos.

1.6.2

Objetivo específico

A partir do objetivo geral, é possível desdobrar vários objetivos específicos para aprofundar a investigação sobre os modelos mentais dos(das) especialistas em IA na criação de algoritmos, sendo os seguintes itens:

- Identificar os principais componentes dos modelos mentais que os(as) especialistas empregam ao projetar algoritmos, focando nas diferenças individuais e coletivas desses modelos;
- Analisar como os(as) especialistas processam informações e tomam decisões durante o desenvolvimento de algoritmos, especialmente em situações que exigem consideração de aspectos sociais e éticos;
- Examinar as estratégias de resolução de problemas utilizadas pelos especialistas em diferentes estágios do desenvolvimento algorítmico, identificando métodos comuns e inovadores aplicados;
- Explorar os critérios e considerações éticas que influenciam as decisões dos desenvolvedores, visando entender como esses fatores são integrados na prática de desenvolvimento;
- Avaliar o impacto dos modelos mentais no design e na funcionalidade dos algoritmos, com especial atenção aos seus efeitos sobre a equidade e inclusão nos produtos finais;
- Discutir diretrizes ou recomendações para o aprimoramento de práticas de desenvolvimento baseadas em uma compreensão mais profunda dos modelos mentais que promovam equidade e eficácia.
- Comparar os modelos mentais entre grupos de especialistas de diferentes gêneros (feminino e masculino), para detectar variações que possam influenciar a abordagem ao desenvolvimento algorítmico.

1.7

Justificativa

O recorte da pesquisa, com a investigação direcionada a especialistas em IA, ou seja, pessoas que desenvolvem em seu dia a dia aplicações com o uso dessa tecnologia, pretende apoiar o aprofundamento das questões sociais mais intrínsecas à realidade em um padrão mais corriqueiro nessa categoria. O panorama do alcance das oportunidades de emprego e formação, as questões de estrutura econômica, modelos familiares, ou seja, toda a existência política dessas pessoas em sociedade, pode refletir na contextualização das decisões de projeto.

Como constatado em (Wang et al., 2022) existem muitos métodos de detecção e mitigação de vieses de gênero propostos na literatura, mas há pouca aplicação generalista desses métodos. Uma vez que muitos desses métodos são fornecidos no contexto de um tipo específico de algoritmo, com tecnologias e parâmetros muito direcionados, essas soluções podem não ser replicáveis uma gama de situações (Wang et al., 2022). Portanto, a intenção de ter um checklist, ou mesmo um framework com conceitos de design universal, pode ajudar a abranger cenários mais diversos.

Dar visibilidade para essas questões, de maneira que estejam localizadas como parte de um framework de projeto, pode levar a discussão para um âmbito mais estrutural. A argumentação é direcionada para haver expansão da consciência e domínio de uso das tecnologias, tornando o indivíduo participativo no desenho de um futuro coletivo mais inclusivo. Personificar os métodos para refletir a conjuntura. Tornar visível as decisões individuais para mensurar os impactos estruturais.

Para a promoção de cenários mais inclusivos precisamos explicar as relações no surgimento dessas condições iniciais. Para decisões mais inclusivas, é preciso ter mais representatividade com espaços preenchidos de maneira igualitária. A tecnologia pode ser caminho e meio para o alcance da equidade de gênero, para falar de uma das desigualdades. Portanto, é relevante o aprofundamento desses mecanismos para o balizamento de políticas públicas que precisam incluir os efeitos da tecnologia em sociedade.

1.8

Metodologia

Quando do surgimento dessa pesquisa, em 2020, os debates isolados sobre a ocorrência de vieses de gênero e raciais na reprodução de contextos por dados por sistemas de IA, começavam a tomar força e ganhar número. Já em 2022/23, houve uma publicação acadêmica bem expressiva na área,

especialmente sobre a comprovação técnica desses desvios com recorte de gênero foram 180 publicações. Logo, para delinear o recorte específico na temática e direcionar o estudo, foi necessário fazer um mapeamento sistemático que envolvesse a tríade: vieses de gênero, projeto de modelagem e algoritmos de recomendação. Essa string foi escolhida com base no exemplo clássico de reprodução de vieses através de um anúncio de vagas em tecnologia no *Facebook* que só atingiu homens através de um sistema de recomendação.

A partir da clusterização dos resultados do mapeamento sistemático, notou-se uma recorrência de relatos técnicos sobre desvios de design e modelagem (projeto), decorrentes da retroalimentação dos dados enviesados. Na categoria "adaptação" estão concentrados estudos que indicam o esforço de adoção de estratégias para mitigação de vieses e a crescente necessidade de adotar a transparência e explicabilidade nos projetos. Desse modo, conectando à fundamentação teórica dessa pesquisa, que traz elementos sobre iniquidade de gênero em ambientes de tecnologia e questões de usabilidade e experiência do usuário, houve o direcionamento para a investigação do modelo mental dos especialistas em IA, a partir da análise de suas tomadas de decisão no desenvolvimento de um algoritmo. Com o apelo do entendimento da parte humana dessa jornada, considerando a conexão usuário-especialista, e, entendendo que o corpo de dados em geral é o direcionador das camadas de aprendizado nos sistemas, é definido que o objeto de estudo é sobretudo a estruturação mental formulada para o desenvolvimento de algoritmos. Para ouvir os especialistas, foram realizadas entrevistas com profissionais do gênero masculino e feminino. A partir do conteúdo transcrito, foram destacadas unidades de contexto (trechos das entrevistas) com declarações que continham os códigos relacionados aos pilares da pesquisa, sendo, usabilidade, IA e experiências pessoais, relacionados à estruturação técnica do projeto dos algoritmos. Por fim foram realizados grupos de foco para discutir direcionadores para um possível checklist de apoio para o desenvolvimento, abordando questões sobre a experiência do usuário, diversidade e acessibilidade. Essa estrutura está esquematizada na figura 1.2

1.9

Estrutura da tese

No panorama atual da sociedade da informação (Gamboa, 1997), a prevalência dos algoritmos como agentes decisórios em múltiplas esferas da vida cotidiana é incontestável.

Alguns passos atrás, para identificar as origens da reprodução das realidades através dos dados, o capítulo 02 aprofunda em temas para cobrir as relações intrínsecas à sociedade de dados. Propomos um entendimento através

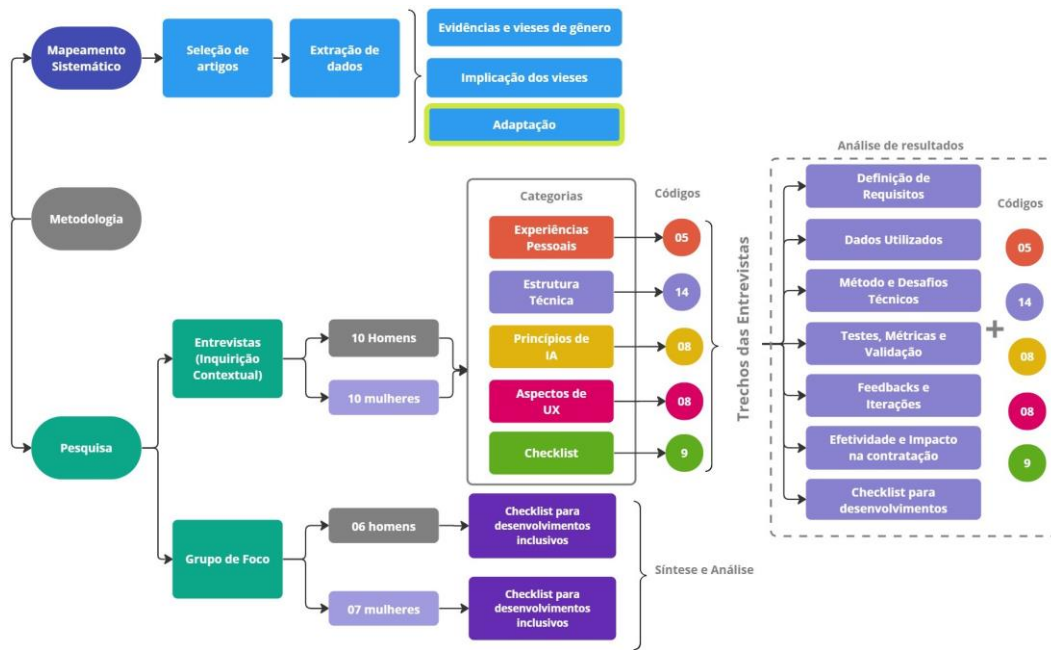


Figura 1.2: Esquemático da metodologia. Fonte: figura elaborada pela autora

das organizações em rede (Castells, 2011). Aqui, a interação digital é escrutinada sob a lente da criação e manutenção de brechas digitais - um fenômeno que reflete as discrepâncias e desigualdades existentes no domínio físico, transpostas e muitas vezes amplificadas para o ambiente virtual (Wajcman, 2006).

O capítulo 3 representa uma parte mais analítica técnica do estudo, sobre aplicação de algoritmos específicos e seu impacto no desenvolvimento profissional feminino. Este segmento aborda a perpetuação da desigualdade de gênero por meio de mecanismos supostamente neutros e objetivos.

O capítulo 4 de metodologia inicia com um mapeamento sistemático da literatura, com o intuito de realizar uma revisão sobre a discussão da ocorrência de vieses em algoritmos de gênero, com foco nas decisões de projeto. Esta investigação inicial pretende dar visibilidade das associações entre o mapeamento das causas de vieses com enfoque em desigualdade de gênero e soluções mitigadoras. Na sequência, há a proposta da coleta de dados através de entrevistas e grupos de foco para a análise dos modelos mentais de especialistas em IA. Há a intenção de entender como esses especialistas estruturam mentalmente o desenvolvimento de algoritmos e se eles perpassam por temas categorizados que representam o envolvimento de aspectos de UX, questões éticas da IA, questões sociais e inclusivas. Nos capítulos 5 e 6 são apresentados os resultados da pesquisa, bem como sua análise.

Este trabalho busca, não apenas destacar a prevalência do viés algorítmico, mas também fomentar um diálogo crítico sobre as potenciais estratégias

para a mitigação de suas consequências. Pretende-se que esse texto impulse reflexões e ações no que concerne à governança algorítmica e suas implicações sociais.

2

Relações digitais e a reprodução da desigualdade de gênero

2.1

A organização em rede e o mapeamento dos algoritmos

A internet com sua estrutura de ligação entre nodos viabilizou novas relações da ordem econômica, cultural e política. Castells (2011) ressalta que a internet como base da tecnologia da informação provê um apoio organizacional para transformações sociais e espaciais; calcadas em fluxos de informação e poder articulam a flexibilização dos sistemas de produção e gestão. Castells (2011) faz uma analogia entre a topologia das redes e as interações sociais, em que a intensidade da frequência entre dois nodos, ou posições sociais, é mais forte se esses nós pertencerem a mesma rede. Essa arquitetura reforça a ideia de grupos e seus determinados privilégios.

A internet, portanto, pode funcionar como um arcabouço organizacional capaz de promover mudanças, criar e até mesmo destruir muitas conexões estruturais da humanidade. Nesse mar de bits, bytes e afins, nós, os usuários, produzimos-consumimos informações e estes estímulos cibernéticos vão sendo transformados em dados. Quantificáveis, manipuláveis e utilizados como moeda de troca, os dados precisam ser colhidos e devidamente encaminhados para as suas “prateleiras” mais adequadas. Aqui entram os algoritmos.

Algoritmos são elementos muito relevantes na estrutura de uso da internet. São equações matemáticas constituídas de mecanismos que atuam nas tomadas de decisão das mais diversas ordens (Buscher, 2016) como recomendações musicais e ofertas de objetos diversos. Porém, a ação algorítmica também gera decisões em campos mais preocupantes. O cálculo do valor de seguro de saúde, em alguns casos, pode ser feito sob as informações filtradas de rastros de dados produzidos no uso da internet. Os indivíduos analisados são considerados aptos ou não para tais serviços, dependendo de suas pontuações. É um ponto crítico e sensível à percepção de cada indivíduo em relação ao poder de influência que esses algoritmos possuem. Será que percebemos a todo momento que somos abordados por esses algoritmos? Conseguimos entender o caminho que esses algoritmos percorrem até aparecerem em nossas decisões?

Para Gillespie (2014) a discussão sobre a utilização dos algoritmos deve

passar por algo além do alcance técnico da equação, precisa atentar para a aplicação do conhecimento humano e o impacto das experiências sociais dos envolvidos em sua produção. A estrutura que perfaz um algoritmo está comprometida funcional e ideologicamente com a produção de sistemas de informação que vão gerar conhecimento e tomar decisões (Gillespie, 2014).

Em termos mais claros, para montar um algoritmo o sujeito detentor da escrita dos códigos define as entradas (inputs) que serão objeto de aprendizado através de informações dos usuários transformadas em dados. As equações vão automaticamente passando por camadas de processamento que o humano não consegue alcançar, nem em conhecimento, nem em termos práticos. Os algoritmos devolvem saídas (outputs) às questões modeladas a princípio (Russell et al., 1995). Chegamos ao conceito de Caixa Preta, que determina esse momento do processamento que foge ao domínio consciente dos criadores das equações. Em *Science in Action* (Latour, 1987) há a descrição que o processo de tornar a técnica invisível acontece para a afirmação de sucesso do próprio processo. Quando uma máquina funciona com eficiência, os acertos ficam acerca das “entradas” e “saídas” e não trata da sua complexidade interna. Desta forma, Latour (1987) ressalta que quanto mais a ciência e a tecnologia obtêm sucesso, mais suas técnicas se tornam obscuras. Em *Introduction to Cybernetic* (Dubberly e Pangaro, 2010) escreve que a teoria da Caixa Preta aborda sobretudo sistemas como objetos reais relacionados aos seus observadores; nessa interação o mote é sobre relacionar quais informações advêm dos objetos e como elas podem ser obtidas e interpretadas.

Esse sistema objeto–sujeito na forma de Caixas Pretas, está inserido em um espaço-tempo mais denso conceituado como cibernética (Wiener, 2019), sendo o estudo interdisciplinar com aplicações tecnocientistas. Peter Galison (1994) narra que Wiener criou o termo cibernética fazendo um paralelo com o funcionamento de um protótipo de mecanismo de defesa antiaérea que ele havia construído (Galison, 1994). Após o domínio da estrutura de tal mecanismo que tenciona processos de controle de informação e comunicação em sistemas complexos, o modelo então é expandido para conceituar uma nova ciência: a Cibernética. Ciência que, segundo Galison (1994), abarca o ato da intencionalidade, do aprendizado e questões concernentes à mente humana.

O existir cibernético desses algoritmos em sistemas complexos trazem um manejo de diferentes perspectivas, um existir em movimento. Para Bucher (2018), os algoritmos devem ser pensados não por uma coisa estática, mas por algo que está se tornando, está se modificando, que não só geram ações, mas também sofrem interferências constantes, estão em constante devir (Jurno et al., 2018). O ato de construir um algoritmo, é elástico assim como seus

resultados; passível de diferentes visões e interpretações. Essa forma elástica transborda em uma interface com possibilidade de uma interação mais humana, mais adaptativa e sensível a intervenções projetuais. Já que é fruto de projeto e está em estado de modificação, a quem interessa essas programações? Qual a relação entre o projeto do algoritmo e a efetividade de suas aplicações nas redes?

O conjunto de dados que produzimos, que são rastros de nossas atividades, preferências e escolhas que nos traduzem, são selecionados por algoritmos para nos representar. À medida que fomos essencialmente adotando aplicações computacionais como meio de expressão, sujeitamos pensamentos e formas de expressão do conhecimento humano a um ajuste computacional para serem codificados (Gillespie, 2014). Portanto, há implicações diretas em validar que esses sistemas automáticos sejam responsáveis por selecionar o que poderá ser mais relevante para disponibilizar ao nosso acesso.

Para Berger e Luckmann (2003), a construção social emerge das relações entre os pensamentos humanos individuais e os contextos sociais que os envolvem, denominado sociologia do conhecimento. As informações, que direcionam o conhecimento, vêm carregadas de significados e se dão no encontro entre os usuários e as fontes de informações (Gamboa, 1997) e em termos práticos para a realidade atual, chegam através de tecnologias. Esse ponto de contato, potencializa as possibilidades de integração e formação cidadã. Pensando que nessa engrenagem dessas ferramentas tecnológicas, os algoritmos se tornam parte da tecnologia de comunicação, passam a ser “instrumentos científicos de uma sociedade em geral” (Gitelman, 2006) e, mais uma vez, são nódulos estruturantes da maneira que entendemos e absorvemos conhecimento em um contexto de vida pública formatados em uma lógica computacional.

Tarleton Gillespie (2014) nomeia esses algoritmos como “Algoritmos de relevância pública”, que conforme descrito anteriormente, possuem o poder de escolher o que nos será apresentado enquanto usuários produtores e consumidores de dados digitais. De maneira, que esse é um mecanismo de produção e certificação de conhecimento.

Todavia, esses algoritmos que avaliam a informação, reproduzem a lógica do conhecimento, formatando o que de relevante compõe determinado significante. Ao incorporar essa estrutura para a construção do conhecimento, delegamos aos algoritmos a decisão do que precisamos saber. Os algoritmos são direcionadores individuais e em conjunto produzem o ecossistema da informação.

Nesse emaranhado de conexões, surgem sombras culturais e informacionais em que pesa entender se o desdobramento dos conteúdos selecionados para

a prática do conhecimento humano pode incorrer em ramificações políticas, no que tange questões sociodemográficas, culturais (Striphas, 2010). No ensaio “A Relevância dos Algoritmos”, Tarleton Gillespie destaca seis dimensões políticas que emergem dessas relações:

1- Padrões de inclusão: o desenho das escolhas por trás do que é incluído em um índice, considerando o que é excluído e como os dados são preparados para os algoritmos.

2- Ciclos de antecipação: ao realizar previsões sobre os usuários, os algoritmos podem antecipar conclusões que podem gerar desvios.

3- Avaliação de relevância: conforme os algoritmos decidem o que é importante para os usuários, e, uma vez que os mecanismos de decisão são opacos, assume-se que pode haver inferências com viés político.

4- A promessa da objetividade algorítmica: a maneira como o caráter técnico do algoritmo é apresentado como garantia de imparcialidade e como essa alegação é mantida diante da controvérsia.

5- Entrelaçamento com a prática: como os usuários adaptam suas práticas para se adequar aos algoritmos dos quais dependem e como podem transformar os algoritmos em territórios de disputa política, às vezes até para interrogar a política do próprio algoritmo.

6 - A produção de públicos calculados: como os algoritmos moldam como os públicos percebem a si por meio da apresentação de informações e como isso pode afetar quem se beneficia desse conhecimento.

Essas dimensões, ou o estágio em que elas se apresentam no momento, podem ajudar a balizar e dimensionar os impactos sociais que os algoritmos e seus vieses podem promover. Algoritmos como ferramentas de conhecimento influenciam e direcionam discursos públicos. E no cerne dessa afirmativa observa-se a capacidade da tecnologia, seus avanços e características; esse dispositivo central não deve ser considerado como algo frio em que suas potências tecnológicas estão isoladas e limitadas por suas capacidades técnicas. Por trás desses artefatos estão pessoas e instituições produzindo recortes e montando um molde para que os resultados de desejo sejam alcançados.

2.2

A tecnologia da informação em tempos de economia da atenção

Castells (1996) diz que ao observar as relações através da internet podemos perceber que a nova economia está organizada em torno das redes globais de capital, gestão e informação; quem detém o saber fazer tecnológico estará competitivamente em vantagem. As funções e processos dominantes na era da informação estão definitivamente organizadas em torno das redes,

Castells (1996) fala em capitalismo informacional. Nossa sociedade se constitui da morfologia social das redes e a difusão de sua lógica sustentável modifica a operação dos processos de produção, experiência de poder e cultura. Embora as formas de organização sociais existam em outra dimensão de tempo e espaço fora das redes, o paradigma da tecnologia da informação traz a base para a expansão pervasiva que atua nas estruturas sociais como um todo.

Com o campo mercadológico estabelecido nas redes, a informação datificada – processo de transformação da informação gerada nos meios tecnológicos em dados (Carpo, 2019) - é ouro.

Zuboff (2019) fala em Capitalismo de Vigilância para esse contexto das redes em que a experiência humana é reivindicada como matéria-prima para a datificação comportamental. Dados são utilizados para incrementar produtos e serviços digitais. O excedente comportamental privado/pessoal alimenta a “maquina algorítmica inteligente” que dá saída a predição de produtos, sendo a própria previsão de comportamentos. Esses produtos são comercializados no campo do “mercado comportamental futuro” (Zuboff, 2019).

Na batalha pela dominação de mercado e maximização dos lucros, o Capitalismo de Vigilância é sedento por ampliar suas fontes de predições através da apropriação dos excedentes comportamentais. Zuboff (2019) define essa caça como extrativismo de dados em que a lógica é de estímulo de produção em larga escala (Zuboff, 2019). Usuários das redes são seduzidos a produzir excedentes através de engajamentos criados especialmente para cada indivíduo ciberneticamente existente. Um mundo único e exclusivo, que estimula a interagir, comprar, escrever, compartilhar, expressar, tudo parecendo grátis, mas convertido em matéria-prima para os produtos manufaturados que serão ofertados no próximo momento oportuno. É necessário capturar as atenções.

Nesse âmbito, como a atenção que cada pessoa dispensa é limitada, há uma disputa econômica para a sua conquista, um recurso escasso e por isso muito valioso. A economia da atenção, um termo conceituado a partir do behaviorismo, é imprescindível para que o Capitalismo de Vigilância se retroalimente e consiga prever e modificar comportamentos.

Zuboff (2019) diz que a economia da atenção é baseada em um axioma onde os excedentes comportamentais devem ser vastos, mas também de natureza diversa. Portanto, o extrativismo de dados não se restringe só aos meios online mas também à vida offline, nos momentos em que as pessoas estão menos alertas para o fato de estarem sendo vigiadas. Conversas durante refeições, interações diárias em comunidade, estão sendo ouvidas, gravadas e datificadas.

A economia da atenção nasce da ideia de que as predições dos comportamentos, ativos altamente lucrativos, devem ser extraídos por padrões incrusta-

dos no self (Zuboff, 2019). Todas as escolhas de compras, o desenho dos rostos, entonação de voz, quantidade de trocas de mensagens são dados escaneados, processados para gerar um self digital. Esses padrões vão sendo criados e impelidos a serem modificados com direcionamentos específicos, caros a quem os compra. E quem são os patrocinadores?

Ao passo que tomamos como ponto um sistema objeto-sujeito que está em constante modificação, para atuar nessa força viva é necessário que o projetista (aquele que escreve códigos) consiga vislumbrar os impactos de sua aplicação no contexto de sociedade. Herbert Simon (1996) em *The Sciences of the Artificial* ressalta que para entender os sistemas, há que se engajar em sua construção e observar como se comportam (Simon, 1996). Lancemos mão de uma crença mais romântica em que há uma força combativa que pode emergir durante o projeto matemático dessas equações algorítmicas. É possível o objeto ser modificado por um projetista holístico (àquele que importa o impacto ao observador). Do ponto de vista em que o design traz ferramentas que aproximam o usuário da criação do projeto, encara-se a situação de intervenção possível no “planejar-fazer” dos sistemas. Seria valioso incutir na programação códigos que refletissem as reais intenções de modelagem para criar mecanismos que deixem os processos menos opacos. A ciência do processo deveria ser passível de registro.

Para Findelli (2000) a abordagem da ciência deve ser a situada ou embutida e não aplicada, considerando que os inquéritos científicos são inerentes à prática do design e estão ambos em movimento (Findeli, 2001). O papel do(a) designer, portanto, é entender a dinâmica morfológica do sistema, a sua inteligência; deve agir no sistema e não sobre ele, designer e usuário são partes pertencentes. Maguire (2001) define que a evolução da prática do design que vai além da prática da ergonomia e fatores humanos, necessita sobretudo de um entendimento mais acurado sobre o contexto de uso, seu ambiente e os atores envolvidos (Maguire, 2001). A situação dos usuários deve ser analisada de maneira momentânea e atualizada, pois as interações são fluídas.

Deste modo em uma situação de projeto: o(a) projetista deve se assegurar que há uma representação adequada do conteúdo, da estrutura, da dinâmica evolutiva e das tendências do sistema. As Caixas Pretas devem ser transpassadas. Segundo Findelli (2001), o estado de resposta de projeto é transitório e pertence a um processo dinâmico, não a uma solução; a produção de um objeto material não é a única resposta as questões de design. As essências das relações são invisíveis e complexas e para fazer parte do processo da busca por solução é preciso trabalhar a inteligência visual: “a chave da nossa era é ver tudo pela perspectiva das relações” (Findeli, 2001). O futuro da Inteli-

gência Visual deve concernir o mundo invisível da consciência humana como pensamentos, motivações, propósito, medo, necessidades, aspirações, etc. Esse mesmo insumo que é datificado com fins mercadológicos, também pode servir como matéria-prima para modificações sociais.

A problematização deve fazer parte do processo bem como o produto/solução também, abrangendo a complexidade do impacto do projeto. Designers fazem parte desse processo, inclusive como cidadãos, segundo Findelli (2000). De acordo com Krippendorff (2004) as atividades multidisciplinares visam maior o esclarecimento de propósitos e significados, considerando que o Design em si como uma atividade pragmática de abordagem empírica que intenciona dar sentido ao mundo que nos rodeia (Krippendorff*, 2004). Há responsabilidade no projeto e também no uso. O usuário consciente faz reverberar a ideia de inteligência coletiva de Pierre Lévy (2003) que enaltece a mobilização efetiva das competências e habilidades que estão presentes nos indivíduos; devem ser coordenadas e usadas a favor da coletividade e esta organização irá acontecer justamente na utilização bem direcionada das tecnologias de informação (LÉVY, 2003). O fluxo de criação, portanto, passa a ser outro: os usuários não são vistos como foco da pesquisa e objetos de insights tão somente, mas passam a fazer parte do processo de ação de criação (Sanders e Stappers, 2008).

2.3

Paralelo digital: consciência algorítmica e a participação democrática nas redes

Muitos serviços rapidamente aderiram aos mecanismos de Inteligência Artificial (IA), desde o apoio a busca e contratação de funcionários, determinação de demissões, indicação de perfis aptos a empréstimos, ou até mesmo o cálculo do tempo que uma pessoa condenada deverá ficar detenta. Muitas das decisões que tradicionalmente eram tomadas por humanos, estão rapidamente sendo feitas por algoritmos (O'neil, 2017).

Porém, a enorme quantidade de dados gerados resulta em formação de personas generalistas pouco sensíveis às particularidades, limitações e privilégios específicos. É mais barato e certo, correlacionar características comuns a maioria dos usuários e traçar perfis a partir de linhas gerais (Invisível, 2020).

Apesar da capilaridade dos sistemas com automação algorítmica em inúmeros países, ainda é pouco estudado o nível do conhecimento, ou mesmo a consciência por parte do grande público sobre a influência e os impactos dos algoritmos em seus cotidianos (Bucher, 2012). A considerar a diversidade

de acesso à infraestrutura, habilidades e padrões de usuários, a área de concentração de estudos denominada Divisão/Exclusão Digital (Digital Divide Studies) faz correlações entre esses índices que podem acarretar o agravamento de desigualdades -(Dutton e Reisdorf, 2019).

Com uma abordagem sociológica (Lutz, 2019), os estudos em Divisão Digital associam questões demográficas a questões culturais como crenças e peculiaridades em relação ao uso da internet, traçando diferentes perfis de “cultura de internet” que se encaixam e/ou produzem experiências com exclusões digitais (Dutton e Reisdorf, 2019). Essas diferenças afetam diversas pessoas em esferas fundamentais de uma maneira não igualitária.

Estar ciente das funcionalidades e impactos dos algoritmos durante a interação em plataformas digitais, redes sociais e sites de busca representa a possibilidade de ter escolhas mais conscientes e críticas. Esta é uma potente habilidade digital que emancipa os usuários e que também demonstra o agravamento da desigualdade de oportunidades (Bucher, 2012). A necessidade de o usuário ter mais acesso aos insights algorítmicos e entender seu impacto informacional é abordada na classificação do conceito de “algoritmo crítico” (Beer, 2019).

A lacuna digital entre estar ciente ou não sobre operação dos algoritmos agrava com o amplo desenvolvimento de aplicações com Machine Learning. Essas estruturas inteligentes que dependem da entrada constante de dados torna os usuários corresponsáveis, que por sua vez, produzem conteúdos inerentes as suas realidades, por vezes restritas. Deste modo, a falta de consciência algorítmica não somente gera uma ameaça a participação democrática nas redes, como também envolve ativamente os usuários na formulação de suas próprias condições de acesso à informação (Bucher, 2012).

A questão é como essa corresponsabilidade, que é amplamente velada, afeta o modo como pensamos nossa participação como usuários atuantes nas redes. Essa infraestrutura inteligente, pesa demandas sobre os usuários transformando-os em “prosumers” - aqueles que tanto produzem quanto consomem (Bucher, 2012). Até que ponto os usuários serão responsabilizados pela construção e fomento das infraestruturas geradas a partir de suas interações?

À primeira vista, o termo “inteligência artificial” sugere uma equivalência com a inteligência humana, talvez surja daí a ideia de corresponsabilidade. Mas voltemos ao início, ao princípio da substituição, onde decisões antes feitas por humanos passam a ser tomadas por algoritmos. A quem serão atribuídas as responsabilidades sobre as decisões automáticas? Já que essas ações serão feitas por máquinas, logo a recorrência dessas operações ganhará um enorme poder de escala. Por sua vez, os impactos dessa tecnologia também ganharão

proporções impensadas?

Muitas instituições e grupos de pesquisa, como HAI – *Stanford University Human-Centered Artificial Intelligence* e *Beijing Academy of AI*, vem reunindo esforços para pensar as questões éticas que envolvem o emprego de IA e suas diversas aplicações. Alguns princípios como privacidade, transparência, explicabilidade, promoção de igualdade e dos direitos humanos são elementos que devem ser correlacionados e assegurados no contexto de criação dos algoritmos e suas respectivas aplicações (Fjeld et al., 2020). A maioria dos documentos, menciona ainda a necessidade de trazer clareza no ato de responsabilização nas tomadas de decisão, tanto no planejamento, quando na continuidade e finalidade dos sistemas. Nesse guarda-chuva da responsabilização há pontos em comum que emergem como guia: a recomendação da adoção e fomento de novas regulamentações, a viabilização de avaliações de impactos sociais, a garantia da reprodução e verificação das aplicações, a responsabilização ambiental e a possibilidade de ser auditável e passível de enquadramento nas leis civis (Fjeld et al., 2020).

Destaca-se o princípio de avaliação de impactos direcionado ao contexto de atuação democrática trazido até aqui. Este princípio trata dos apontamentos específicos de questões de direitos humanos e da necessidade de identificar previamente e mitigar os possíveis impactos negativos das tecnologias de IA (Access Now, 2020). Um caminho para medir esses impactos é avaliar os riscos e oportunidades para o fomento dos direitos humanos. Segundo o relatório Access Now (2020), é recomendado que as organizações de referência sejam sempre consultadas, como grupos que se organizam em prol da segurança e representatividade dos direitos de grupos minoritários; para a prevenção de danos devem ser abordadas questões de ordem emocionais, sociais, ambientais, comunicacionais e outros temas de ordem não-econômica. Há muito o que ser discutido sobre a ética e informação no uso de sistemas automáticos, considerando a elasticidade das aplicações.

2.4

Brechas Digitais: reprodução de desigualdades e o outro digital

A descentralização da informação nas redes (Castells, 2011) ramificou estratégias de comunicação e mobilização de grupos. A capacidade de mobilizar expressões individuais é impulsionada pela quebra das paredes entre o que é público e o que é privado (Costa, 2018), seres presentes nas redes ou *smart people* funcionam como sensores que produzem e consomem dados constantemente (Invisível, 2020).

Essa facilidade generalizada de conexões e produção de conteúdo nas re-

des, gera uma variedade cultural e social reproduzida digitalmente. Aliado ao conceito de “Divisão Digital”, todas as questões que diferenciam as oportunidades de acesso podem dar brechas ao acontecimento do “Outro Digital”. Gordano (2009) faz uso do termo “Outro Digital” para representar aqueles indivíduos subvalorizados e não pertencentes ao universo de acesso e apropriação da tecnologia da informação. São indivíduos sem material simbólico para ocupar o lugar de usuário modelo ou mesmo para produzirem volume de dados suficientes que os representem (Gordano, 2009).

Além disso, a representação majoritariamente masculina no ato de desenvolvimento dessas ferramentas tecnológicas pode gerar silenciosamente um movimento de retrocesso nas questões de igualdade de gênero. A esse efeito Wajcman (2006) atribui a denominação de brechas digitais (Wajcman, 2006). De acordo Buolamwini (2018) a inteligência das máquinas é construída através do aprendizado gerado pela observação dos dados apresentados. Uma vez que os algoritmos, agentes do processamento inteligente, têm capacidade de interpretar uma grande massa de dados, se esses vierem carregados de conceitos estereotipados, os vieses de gênero sexistas serão perpetuados (Buolamwini e Gebru, 2018), sem falar em outras questões interseccionais com recortes sociais, de classe e raça (Crenshaw, 2013).

Há, portanto, o fenômeno imperativo da “codificação” dos vieses reproduzidos. Uma vez que os algoritmos projetados por humanos inseridos em contextos diversos, consomem dados produzidos por humanos em sociedade, é natural serem reproduzidos vieses, por vezes preconceituosos, a partir dessa leitura virtual da sociedade. Trata-se de dados produzidos simultaneamente, de maneira global e descentralizada (Li e Etchemendy, 2018); o controle sobre a natureza ética do direcionamento dessas aplicações, se torna bem complexo. Observamos a organização social virtual sendo reproduzida a partir de dados “colonizados” e “viciados” (Silva, 2019).

O olhar para os sujeitos envolvidos no sistema, os ditos usuários operários, deve ser cuidadosamente localizado de dentro para fora. É necessário ser sensível à diversidade. A variedade cultural e social junto às oportunidades de acesso e profundidade de aprendizado podem dar brechas ao acontecimento do “outro digital”. Gordano (2009) faz uso do termo “outro digital” para representar aqueles indivíduos subvalorizados e não pertencentes ao universo de acesso e apropriação da tecnologia da informação. São indivíduos sem material simbólico para ocupar o lugar de usuário modelo. A narrativa hegemônica de dominação cultural recai sobre a imbricação dos usos das redes (Natansohn, 2014). Quem detém o poder de projeto é quem poderá moldar o usuário padrão, que provavelmente reproduzirá uma lógica carregada de vieses dominadores, a

menos que esse não seja o interesse.

2.5

Aspectos práticos para a desigualdade: Segregação Vertical de Gênero

Sob um olhar analítico para diferenças de gênero em âmbitos culturais, trazemos o seguinte questionamento: o que faz homens e mulheres ocuparem papéis tão distantes entre si, e ainda mais desiguais quando fazemos um recorte profissional? Conforme o Relatório do Fórum Econômico Mundial (2018), as mulheres levarão ainda 100 anos para conseguir equidade de direitos e equiparação salarial em comparação aos homens.

Segundo o Censo da Educação Superior (2022) sobre o perfil dos discentes nos cursos de graduação, a maioria das pessoas matriculadas é do sexo feminino. Já em relação aos docentes, ainda segundo a referida pesquisa, a maioria é de homens com idade média de 41 anos.

Consoante o relatório *Education at Glance* (OCDE, 2019) mulheres tem 34% de chance a mais do que os homens em obter seus diplomas. Já em um panorama de mercado, o relatório indica uma taxa de empregabilidade para os homens de 89% em contraponto a 82% para as mulheres.

A diferença se agrava no desenvolvimento de carreiras, pois as mulheres ocupam uma média de 36,67% dos cargos em posições de liderança. O afunilamento de ocupação feminina ao longo do desenvolvimento profissional é denominado “efeito tesoura” (Moss-Racusin et al., 2012).

A esta superposição de representatividade masculina em cargos mais altos, tanto em contextos de mercado de trabalho quanto acadêmicos, atribui-se o termo Segregação Vertical de Gênero, sobretudo na comparação entre a ocupação específica de cargos entre os sexos (Charles e Grusky, 2005). A iniquidade de gênero pode refletir as escolhas das mulheres por carreira em áreas menos valorizadas em termos financeiros, que também passam por questões estruturais e culturais, como a falta de confiança em assumir estratégias para galgar promoções (Cech, 2013).

Em termos de avaliação da produtividade é preciso correlacionar fatores diversos. A quantidade inferior na produção de artigos, estudos e relatórios pode ter causa no acúmulo de tarefas que oneram o tempo em que a profissional deveria dedicar a geração de resultados mais efetivos (England, 2005).

Para esquematizar as causas dessa desigualdade, a pesquisadora Katherine Weisshaar (2017) desenvolveu um modelo de equação com parâmetros menos abstratos para relacionar o reduzido número de mulheres em cargos de níveis mais altos da própria Universidade. O estudo de Weisshaar apresenta indicadores distribuídos por três vieses (Weisshaar, 2017):

1. Especificação das áreas profissionais e contexto de ambiente de trabalho

- Menos opções de cargos para mulheres
- Preconceito na contratação
- Desigualdade nos processos e regras de promoção

2. Produtividade acadêmica

- Preferência de orientação e tutoria à pesquisadores homens
- Natureza dos projetos
- Desigualdade nos processos e regras de promoção
- Jornadas duplas e triplas das mulheres que afeta a produtividade

3. Avaliação geral

- Preconceitos explícitos de gênero
- Formação de redes de fortalecimento às quais mulheres não pertencem, por falta de disponibilidade de tempo ou por exclusão
- Falta de confiança por parte da própria pesquisadora em processos auto avaliativos.

Segundo Weisshaar (2017), a decomposição dos dados apurados indica que a segregação vertical é fomentada em menor parte pela diferença de produtividade. Com maior expressividade, os resultados sugerem que há bastante influência de contextos culturais subjetivos, decorrendo em menores taxas de promoção profissional entre as professoras

Fica claro que ainda há muitos fatores invisíveis e não mensuráveis. Contudo, o esforço na classificação e decomposição dos indícios das causas da iniquidade de gênero é uma importante ferramenta para a trajetória da busca por uma distribuição mais igualitária de oportunidades.

3

A tecnologia e o design de Algoritmo centrado no humano: questões de projeto

3.1

Aspectos técnicos

Algoritmos são recursos que podem ser embutidos em softwares tradicionais trazendo funcionalidades ou fazem parte de sistemas inteligentes que são parte de engrenagens mais abrangentes, como plataformas. Sistemas inteligentes tem incorporados componentes de *Machine Learning* para apoiar suas operações, sendo denominado *ML-enabled systems*. A principal diferença entre os dois sistemas são os dados que, de certa forma, substituem os códigos e direcionam as decisões e comportamento dos sistemas. Portanto, é recomendado que as validações e testes dos modelos de ML sejam minuciosamente executados, conforme a boa prática realizada para os códigos dos softwares non-ML (Arpateg et al., 2018). Os algoritmos seguem comandos com sintaxes estruturadas para permitir a execução das tarefas (Arpateg et al., 2018). São recursos que funcionam nos bastidores e possuem características inteligentes, conseguem ler, ouvir, sentir e de acordo como forem configurados podem trazer experiências únicas para os usuários através do uso de produtos e serviços.

Ter o domínio do fluxo de decisões e a jornada dos dados se faz necessário para a estruturação de um algoritmo, compreendendo todas as etapas entre a entrada do problema e a saída dos resultados. Mais especificamente, essas decisões pousam no campo da ciência de dados. A disciplina conhecida como ciência de dados encontra seu lugar de convergência entre o domínio do conhecimento matemático, estatístico, computacional e de engenharia, junto ao entendimento profundo dos dados específicos de diversas aplicações (Rezende, 2003).

A onipresença dos dados na era digital confere à ciência de dados uma abrangência que se estende a todos os setores, e seu emprego vem crescendo de maneira substancial na resolução de desafios. Com o propósito de conduzir análises que embasem o processo de tomada de decisões, envolvendo geralmente o manuseio de extensos volumes de dados de maneira organizada e sistemática, são realizadas etapas para o desenvolvimento de algoritmos (Kalinowski et al.,

2023).

Segundo Rezende (2003) a ciência de dados é um conjunto de métodos para coleta, armazenamento, processamento, análise e visualização de dados. Seu propósito primordial consiste na construção de modelos a partir da própria informação disponível de maneira geral, esses dados são adquiridos de forma não preestabelecida, sem a realização de experimentos planejados ou a garantia de que a amostra seja representativa. A coleta e o armazenamento de informações, representam uma etapa de extrema relevância dentro desse processo.

A jornada inicia com a escolha de um subconjunto de dados destinado ao refinamento do modelo. A escolha dos dados iniciais, ou a formação do *dataset*, compreende a etapa de pré-processamento que envolvem a escolha, a depuração e a adaptação dos dados. O passo subjacente a essa fase é a formatação do conjunto de dados destinado ao treinamento do modelo, denominado conjunto de treinamento. Esse subconjunto é selecionado como uma amostra confiável, podendo, em algumas situações, equivaler à base de dados integral. Os dados presentes nesse conjunto de treinamento desempenham um papel fundamental no desenvolvimento do modelo. Ao longo do processo de construção do modelo, ocorre uma divisão dos dados para ajuste de parâmetros e validação dos resultados. Geralmente, a etapa de pré-processamento consome entre 60% e 80% do tempo total no processo de mineração de dados e proporciona informações cruciais que orientam tanto a análise quanto o ajuste dos modelos (Rezende, 2003).

Após a fase de pré-processamento, a etapa de modelagem e inferência se resume, em termos simples, a ajustar o modelo de maneira específica e apropriada para abordar a problemática em foco. É justamente a etapa que vamos aprofundar observação com a intenção de elucidar pontos na jornada da modelagem de projeto em que as decisões mais parciais de desenvolvedores podem impactar em resultados menos democráticos ou inclusivos.

Segundo Kalinowski (2023), a fase de modelagem e inferência se desdobra em três etapas distintas: identificar os modelos viáveis e apropriados para cada categoria de problema; elencar os parâmetros que compõem esses modelos, com base nas instâncias e nas variáveis que passaram pelo pré-processamento; avaliar o desempenho de cada modelo, recorrendo a métricas e a um processo justo de comparação.

3.2

Auditando Algoritmos e outras técnicas de transparência

Quando falamos em dar visibilidade a estrutura modelada ou ao projeto do modelo, falamos em transparência algorítmica (Pasquale, 2011). Em relação ao desenvolvimento de sistemas de IA, a transparência deve ocorrer na indicação dos dados, no processo e no software, o que abrange a coleta de dados, o processamento de dados, as escolhas de modelagem e a qualidade do treinamento (Kim e Doshi-Velez, 2021). Não é de hoje a demanda para o entendimento dessas estruturas, seja para que esses dispositivos sejam monetizados, seja para utilizá-los como alçôques comprovados de algum efeito nocivo aparente. Na década de 60, a partir um esforço conjunto entre as empresas norte-americanas de American Airlines e IBM foi criado um software para reservas automatizadas de tickets de viagens aéreas, o SABRE (Semi-Automatic Business Research Environment) (Redmond e Smith, 2000). O SABRE revolucionou a indústria de aviação, reduzindo o tempo necessário para reservar um bilhete de avião de até três horas com papel e telefone para apenas alguns minutos via computador (Redmond e Smith, 2000).

Mediante a análise do sucesso das operações, a American Airlines, entendeu que a ordem em que apresentava os resultados de busca afetava na escolha efetiva dos usuários. A empresa começou a manipular a ordem da listagem de resultados da busca para favorecer a receita das vendas. Este mecanismo denominado “ciência de tela” (Petzinger Jr, 1996) os levou a investigações antitruste pelo congresso americano, mas em sua defesa a empresa questionou se seria antiético usar estratégias para melhorar suas vendas utilizando um algoritmo que eles mesmo tinham criado. Em resposta as autoridades americanas instituíram que os algoritmos deveriam ser transparentes e regulamentaram critérios para a divulgação de voos (Redmond e Smith, 2000).

Uma vez compreendida a necessidade da exposição da modelagem dos algoritmos e trazendo para a realidade atual, é proposta a técnica de auditoria de código (Sandvig et al., 2014), com inspiração na técnica de auditorias tradicionais. Sandvig (2014) propõe cinco técnicas que dispõem de desenhos diferentes das interações das partes interessadas: auditoria de usuário não-invasiva; sock-puppet audit; auditoria colaborativa; auditoria de código; e auditoria de raspagem. Dentre essas opções, a auditoria de código seria a mais recomendada para entender os momentos de modelagem e decisões de projeto, porém a mais difícil de aplicar, visto que muitos códigos são mantidos em sigilo pelas empresas, por serem considerados ativos de propriedade intelectual. O esquema se dá pela investigação na perspectiva de especialistas e pesquisadores do código especificado, com anotações e submersão no projeto escrito, com

interação conforme diagrama na figura 3.1.

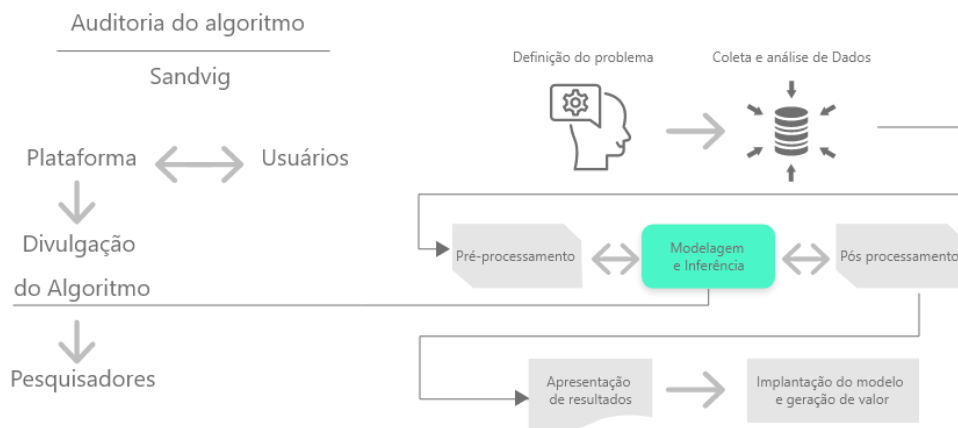


Figura 3.1: Esquemático de auditoria de códigos

O método propõe que especialistas consigam mapear e reproduzir as cadeias de decisões, os parâmetros estabelecidos para escolha de *datasets*, o próprio *dataset* e aplicação de módulos técnicos específicos.

Essa reprodução em termos práticos é complexa e dificultosa, visto que são características muito intrínsecas a serem alocadas. Porém, havendo a possibilidade, esse caminho da apuração pode levar a entender possíveis aparatos criados, muitas vezes não de forma consciente, que trazem desdobramentos ao serem aplicados em sociedade.

Um exemplo de sucesso da aplicação desse método é citado em estudos que envolvem o projeto Gender Shades Buolamwini e Gebru, 2018 realizado pela Algorithmic Justice League. A equipe liderada pela pesquisadora Boulamwini conseguiu analisar detalhes de recursos algorítmicos empregados em três principais tecnologias de mercado para reconhecimento facial e comprovaram que os sistemas acertavam mais na identificação de pessoas brancas e homens. A partir da análise da taxa de erro nas respostas do sistema, a pesquisa conseguiu identificar as causas raízes e atestar que as escolhas feitas sobretudo nas orientações de *dataset* e seleção de recursos causam vieses de reprodução de preconceitos interseccionais, passando por raça, gênero e classe (Buolamwini e Gebru, 2018).

Atualmente, com toda a popularização e aumento de acesso às tecnologias baseadas em IA, há mais ferramentas que têm como objetivo fornecer visões sobre o processo de tomada de decisão dos modelos de IA, aumentar a transparência e possibilitar que os usuários compreendam e confiem nas recomendações ou previsões feitas pelos modelos. Para Kim e Doshi-Velez (2021) as principais ferramentas endereçadas a explorar a transparência nos processos de sistemas com IA Kim e Doshi-Velez, 2021:

1. LIME (Explicações Locais Interpretáveis para Modelos Agnósticos) é uma ferramenta que oferece explicações para previsões individuais aproximando o comportamento do modelo de forma local.
2. SHAP (Explicações Aditivas de Shapley) é uma ferramenta que atribui valores de importância às características em um modelo, permitindo uma compreensão global da contribuição das características.
3. XAI (Inteligência Artificial Explicável) é um campo que se concentra no desenvolvimento de ferramentas e técnicas para tornar os modelos de IA mais transparentes e interpretáveis.
4. Modelos baseados em regras, como árvores de decisão ou listas de regras, fornecem regras explícitas que podem ser facilmente compreendidas e interpretadas por humanos.

Técnicas agnósticas ao modelo, como análise de importância de características ou gráficos de dependência parcial, podem ser utilizadas para entender o impacto de diferentes características nas previsões do modelo.

Essas ferramentas têm como objetivo fornecer visões sobre o processo de tomada de decisão dos modelos de IA, aumentar a transparência e possibilitar que os usuários compreendam e confiem nas recomendações ou previsões feitas pelos modelos.

3.3

Elementos e princípios da Inteligência Artificial Centrada no Humano: um framework para desenvolvimento de sistemas inteligentes

Uma constante argumentação para que os sistemas inteligentes tomem caminhos não desejados, é a da opacidade nos processos de formação do aprendizado. As entradas e saídas podem ser definidas, porém, o processamento do aprendizado de máquina vai percorrendo camadas e tomando decisões inacessíveis ao conhecimento humano (Verganti et al., 2020). O desafio é criar componentes que atuem nessa operação automática para contextualizar as diferentes realidades e áreas de conhecimento, aproximando especialistas às necessidades dos usuários. Não há espaço somente para profissionais de tecnologia, mas é necessário o envolvimento de profissionais de sociologia, filosofia, direito, geografia, psicologia, etnografia e muitas outras especialidades.

Pensando no domínio dos pontos visíveis da modelagem de dados, para além das caixas pretas, algumas Universidades e Centros de Pesquisa se uniram e estabeleceram o conceito HAI – Human Centered AI (Wei Xu, 2019). É instituído pensar as fronteiras da IA não somente guiada pela tecnologia, mas

sobretudo por questões humanitárias e éticas-sociais, pensando no contexto das experiências dos usuários. Segundo Wei Xu (2019), aplicando as diretrizes de HAI, são mapeados os seguintes objetivos principais: as questões tecnológicas que implicam o desenvolvimento da IA devem refletir a profundidade e nuances da inteligência humana; a IA deve aprimorar as capacidades humanas e não almejar sua substituição; e por fim é necessário manter o foco no impacto que a aplicação de IA pode causar na vida humana.

Considerando que IA é uma automação de tomadas de decisão e que devido à natureza da tecnologia o processo de decisão não é intuitivo e nem tampouco transparente (Gillespie, 2020), os usuários podem ser tomados por perguntas sem resposta ao utilizar, ser influenciado ou direcionado por sistemas automatizados. Esses momentos opacos acarretam falta de confiança, nas decisões de adoção dos sistemas e, portanto, na eficiência das entregas dos sistemas (Bucher, 2016).

Segundo Li e Etchemendy (2018) ao longo do desenvolvimento da IA houve um grande descompasso entre as evoluções tecnológicas e o acompanhamento das necessidades humanas e por sua vez a ineficiência da absorção da tecnologia por usuários que no geral não foram ouvidos e não são atendidos. É recente as adequações mais humanas da tecnologia, somente na terceira onda - 2006 em diante - a finalidade das aplicações deixa de servir somente às evoluções tecnológicas, mas também atendem a questões centradas no humano, gerando, portanto, experiências positivas (Wei Xu, 2019).

Olhando para essa lacuna e entendendo que para operar nos pontos visíveis da construção dos sistemas e para promover uma boa experiência ao usuário a tecnologia deve ser compreensível, explicável, útil e com boa usabilidade, Wei Xu(2019) propõe um framework com base na HAI (human- centered AI). Além das questões éticas para garantir segurança e prática de direitos humanos, a tecnologia deve abarcar de maneira sensível à profundidade das reflexões humanas e por fim, deverá ter em sua construção o envolvimento de métodos de Ergonomia para assegurar ao usuário uma boa usabilidade (Wei Xu, 2019).

O conceito de Inteligência Artificial Explicável (Explainable AI) compreende o emprego de um conjunto de processos e métodos para tornar modelos de IA transparentes, evidentes e claros, não só nos pontos de contato visuais (interface) mas também na relação entre os insumos de entrada e saídas dos sistemas, sendo esses objetos comunicados sobre suas existências e funções (Nori, 2019).

As resultantes desse contexto podem ser a esquematização visual dos modelos e arquitetura das soluções, interfaces adaptativas, diálogos o mais

próximo do comportamento natural humano, com mediação humana e observação com mecanismos interdisciplinares, característica de profissionais de IHC (Interface Humano-Computador). Profissionais estes, que devem conduzir as experimentações com rigorosos métodos para conseguirem mensurar os protocolos de pesquisa (Verganti et al., 2020).

Ainda sobre a corresponsabilização compulsória dos usuários, a classificação dessa interação como *human-machine teaming* compreende o conceito de uma formação de uma equipe entre o humano e a máquina (Xu, 2019). Assumindo que essa relação tem dupla validação e aproveitamento, ambas as forças devem ser valorizadas, bem como a capacidade de produção, pensamento e transformação dos dados. Os humanos são parte fundamental nessa engrenagem e não somente consumidores-usuários que se restringem a interações pontuais com o sistema.

Considerar os humanos como parte do sistema e se ocupar da qualidade dessa interação é falar sobre a Experiência do Usuário (UX – User Experience). UX pode ser definida como “sentimento momentâneo na interação com um produto ou serviço, sobretudo sob a ótica da avaliação bom-ruim” (Hassenzahl, 2008). O campo de estudo de UX se interessa, portanto, pela subjetividade do uso do artefato, que acontece de forma dinâmica, fenomenologicamente orientado ao presente e passível de constantes mudanças (Hassenzahl, 2008).

A partir dessa definição, algumas perguntas são decorrentes, portanto, acerca da prática de UX: Como entrar na cabeça das pessoas enquanto elas interagem com os produtos? Como rastrear UX ao passar do tempo? Como avaliar as interações momentâneas? Como integrar sentimentos momentâneos e atitudes posteriores a interação à noção de acúmulo de sentimentos? E como questionamento central para este estudo, como compreender a percepção das pessoas ao interagir com sistemas baseados em algoritmos com aprendizagem de máquina?

Para que esses questionamentos sejam atendidos se torna imperativo o papel dos profissionais de UX no ato de modelagem dos cenários para a construção do algoritmo. Deve haver uma inquirição contextual para desenhar as personas e suas necessidades e particularidades das interações localizadas nos ambientes específicos. As especificações dos nichos e o foco no usuário devem traduzir os caminhos para a interação com os sistemas em linguagem fácil, adequada, com comunicação compreensível com foco na boa usabilidade. Esse processo requer maturação dos dados, com prática de uma visão sistêmica sob os resultados gerados a partir do uso dos métodos mais adequados, testes de usabilidade e validação dos produtos (Hassenzahl, 2008).

Para explorar, comunicar e entender os resultados, é necessário um

processo bem estruturado. Segundo Heer e Shneiderman (2012) a escolha do conjunto de dados é uma parte fundamental para garantir o valor dos parâmetros que serão trabalhados no desenvolvimento dos algoritmos. O valor de um parâmetro é trabalhado em sua visualização por cores, formas e dados plotados. É necessário observar vínculos na natureza dos parâmetros dos conjuntos escolhidos para demonstrar a relação dos resultados.

3.4

Questões para o Design do modelo de IA design de algoritmo centrado no humano

Acompanhando a evolução da tecnologia e observando questões sociais geradas a partir da interação com a tecnologia em sociedade, a partir de 2010 as publicações e pesquisas em IHC foram agregando, para além de constatações estatísticas, novas abordagens focadas em colaborações, conexões, emoções e comunicação. A natureza das interações varia conforme os artefatos, englobando computação gestual, reprodução de linguagens naturais, dispositivos vestíveis, sustentabilidade, Big Data, acessibilidade e outros tópicos (Liu et al., 2014). Portanto, os tópicos tratados em pesquisas de IHC continuam mudando ao passo do desenvolvimento tecnológico, necessidades sociais, patrocínios governamentais (políticas públicas) e frustrações dos usuários (Lazar; Feng; Hochheiser, 2017).

Para atualizar e fomentar o ensino e prática de HCD (Human Centered Design), Lazar, Feng e Hochheiser (2017) fazem um mapeamento de toda a atuação projetual da área e apontam as três características fundamentais que envolvem projetos com HCD – atuação sistêmica, orientação por projeto e foco no equilíbrio na relação entre desempenho e bem-estar. Dul et al. (2012) identificam quatro principais grupos de interessados em um sistema de design com HCD:

1. **Sistema de atores:** funcionários, usuários, aqueles que fazem parte direta ou indiretamente dos processos e tem seus desempenhos influenciados diretamente;
2. **Sistema de especialistas:** profissionais especializados no tema de interesse e especialistas em ergonomia (HCD);
3. **Sistema de tomadores de decisão:** gerentes, contratantes, responsáveis pela implementação do sistema;
4. **Sistema de influenciadores:** mídia, governantes, organizações regulatórias, cidadãos que tem atuações públicas, em geral, com interesse no sistema de trabalho em questão.

Fazendo um paralelo com o desenvolvimento de sistemas inteligentes e projetando todo o contexto até aqui discutido, podemos espelhar a partir desse levantamento e fazer um esboço relacionando os principais interessados que devem ser considerados no desenvolvimento de algoritmos centrados no humano com mais garantias de boas experiências de uso e promoção de igualdade, preocupação ética e sustentável.

1. **Sistema de atores com corresponsabilidade:** usuários, produtores de dados, aqueles que fazem parte direta ou indiretamente dos processos e tem seus desempenhos influenciados diretamente;
2. **Sistema de especialistas com função reguladora e de desenvolvimento:** desenvolvedores, especialistas no tema de interesse, representantes de minorias, sociólogos, psicólogos, historicistas, profissionais que representem a diversidade interseccional;
3. **Sistema de tomadores de decisão responsáveis pela viabilização de infraestrutura de desenvolvimento e promoção de bem-estar:** gerentes, contratantes, responsáveis pela implementação do sistema e empresas envolvidas no consumo de dados;
4. **Sistema de influenciadores promotores da conscientização do alcance do uso da tecnologia:** mídia, governantes, organizações regulatórias, cidadãos que tem atuações públicas, em geral, com interesse no sistema de trabalho em questão.

Essa relação dos agentes, suas interfaces e dinâmica de abrangência pode ser analisada na figura 3.2

Maguire (2001) define que a evolução da prática do design vai além da prática da ergonomia e fatores humanos, ela necessita sobretudo de um entendimento mais acurado sobre o contexto de uso, seu ambiente e os atores envolvidos. A situação dos usuários deve ser analisada de maneira momentânea e atualizada, pois as interações são fluídas. Já para Krippendorff (2004) as atividades multidisciplinares procuram o maior esclarecimento de propósitos e significados, considerando que o design em si é uma atividade pragmática de abordagem empírica, intencionando dar sentido ao mundo que nos rodeia.

Numa abordagem coparticipativa, o fluxo de criação, porém, passa a ser outro: os usuários não são vistos como foco da pesquisa e objetos de insights tão somente, mas passam a fazer parte do processo de ação de criação (Sanders e Stappers, 2008).

Os algoritmos, portanto, devem ser centrados nos humanos e não apenas baseados em dados sem contextualização. O entendimento do contexto que as

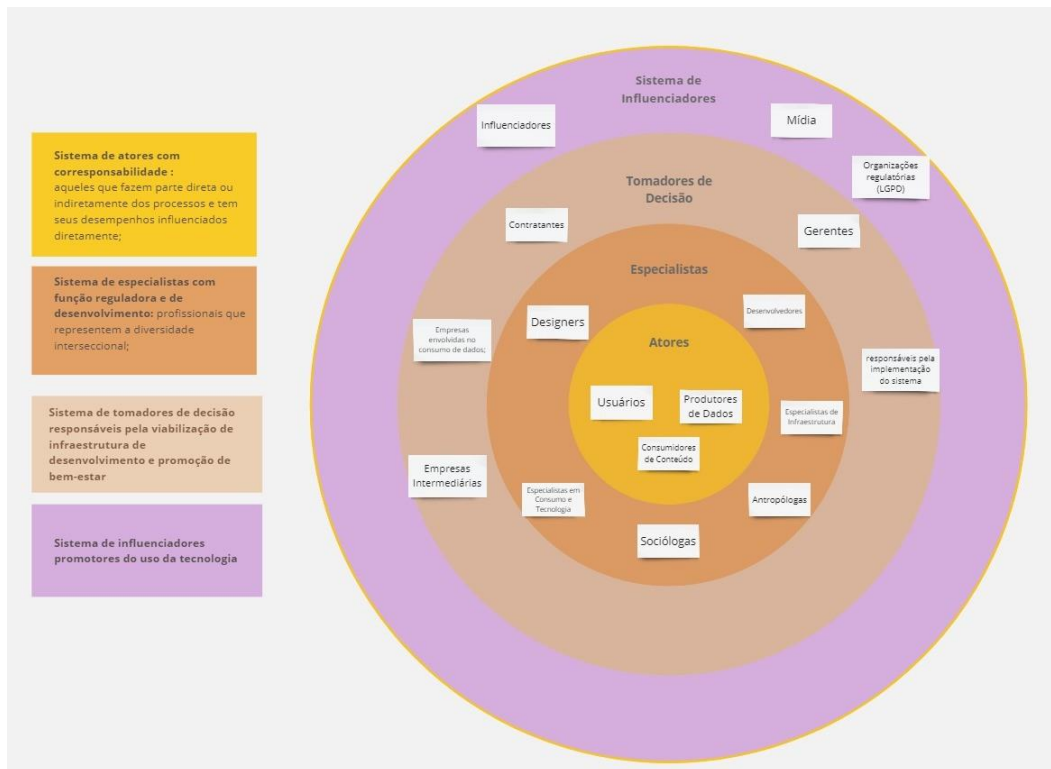


Figura 3.2: Ecosistema de Agentes para desenvolvimento de sistemas inteligentes com técnicas de HCD

peças estão inseridas, não deve ser resultado apenas do consumo de dados, mas estes precisam ser correlacionados e modelados com pesos e medidas especializadas, que façam sentido considerando as particularidades, desejos e necessidades geradas pelo contexto. Pensando esse sistema de coparticipação, a forma de comunicar, alterar, direcionar e expressar através da interface são pontos essenciais, que precisam ser cuidados por especialistas e interessados em garantir uma boa usabilidade consciente.

3.5

Design inclusivo, Design Universal e Design para todos

É possível haver igualdade de acesso, mas não existe igualdade na capacidade de utilização (Wolton, 1999). Essa observação, referente a um conceito embutido em Design Universal, vai ao encontro do contexto das brechas digitais (Gordano, 2009) que se refere não apenas à falta de acesso tecnológico, mas também às disparidades no conhecimento e habilidade para efetivamente utilizar essas tecnologias. Desta forma, convergindo para melhorar as condições de acesso, diferentes abordagens do design como, Design Inclusivo, Design Universal e Design para todos, se concentram principalmente em aumentar a acessibilidade de sistemas interativos para a maior gama possível

de usuários. Reservando algumas particularidades nos conceitos, vamos focar na contextualização de Design Universal.

Pesquisas mais recentes sugerem que o termo 'design universal' pode ser utilizado de maneira intercambiável com 'design para todos' (Stephanidis, 2001). Design para todos é uma abordagem ética que exige um grau de flexibilidade, considerando que o design para a maioria, ou seja, para mais pessoas, precisa também incorporar o design para a minoria, ou seja, para menos pessoas (Bougie, 2002). Desenvolvido por Ronald L. Mace, arquiteto, designer de produtos e educador, o Design Universal é um conceito que busca atender às necessidades das pessoas, sem distinção de idade, habilidade ou status social (Mace, 1996). Esse conceito evoluiu dos princípios de design sem barreiras e design acessível. Mace destaca que o que pode ser acessível para uma pessoa pode representar uma barreira para outra, e ressalta a complexidade do design, enfatizando que remover barreiras é insuficiente sem uma abordagem mais ampla.

Na perspectiva atual, destaca-se a relevância de incorporar o Design Universal em um contexto baseado no princípio da diversidade e heterogeneidade dos usuários e dos cenários de uso. Esta abordagem requer uma elasticidade quanto à definição de "usuário" e "contextos de uso". Sob esta ótica, o "usuário mediano" é visto como um modelo teórico que não reflete as variadas diferenças individuais e contextuais que são cada vez mais prevalentes no cenário social contemporâneo. Desta forma, como sugerido por Bougie em 2002, é essencial integrar estudos sobre as particularidades das minorias nas metodologias tradicionalmente aplicadas às majorias, para expandir o leque de atributos individuais, contextuais e organizacionais que definem os usuários (Almeida, 2015).

Por definição, Design Universal é a criação de produtos e ambientes que sejam utilizáveis por todas as pessoas, na maior medida possível, sem necessidade de adaptações ou designs especializados (Connell, s.d.). A implementação do Design Universal nos leva a considerar não apenas a amplitude do acesso, mas também a igualdade na capacidade de utilização. As desigualdades de conhecimento não se modificam com a simplificação do acesso e da utilização (Wolton, 1999). A essência do Design Universal, portanto, enfatiza a importância de incorporar ajustes para uma variedade de usuários já nas fases iniciais de desenvolvimento de produtos e aplicações, minimizando a necessidade de adaptações futuras. Essa abordagem assegura que as aplicações sejam acessíveis a um espectro amplo de usuários (Zahm, 1997). Desta forma, a implementação de métodos adequados é necessária para viabilizar a personalização através da seleção de componentes de interface alternativos, adaptando-se às necessidades

específicas de cada usuário (Stephanidis, 2001).

Para detalhar o conceito, foram estabelecidos sete princípios de design universal (Connell, s.d.) :

1. Uso equitativo: o design é útil e comercializável para pessoas com habilidades diversas.
2. Flexibilidade de uso: o design acomoda uma ampla gama de preferências e habilidades individuais.
3. Uso Simples e intuitivo: o uso do design é fácil de entender, independentemente da experiência, conhecimento, habilidades linguísticas ou nível de concentração atual do usuário.
4. Informação perceptível: o design comunica as informações necessárias de forma eficaz ao usuário, independentemente das condições ambientais ou das capacidades sensoriais do usuário.
5. Tolerância ao erro: o design minimiza perigos e as consequências adversas de ações acidentais ou não intencionais.
6. Esforço físico reduzido: o design pode ser usado de maneira eficiente e confortável, com um mínimo de fadiga.
7. Tamanho e espaço para aproximação e uso: são fornecidos tamanho e espaço adequados para aproximação, alcance, manipulação e uso, independentemente do tamanho do corpo, postura ou mobilidade do usuário.

Observações finais

À medida que concluímos a discussão sobre a influência crítica dos algoritmos e o papel da ciência de dados em sistemas interativos, reconhecemos a complexidade e os desafios inerentes ao desenvolvimento dessas tecnologias. O progresso feito no campo da inteligência artificial eleva a necessidade contínua de abordagens metodológicas robustas para garantir que essas ferramentas sejam tão inclusivas e equitativas quanto pretendido. Desta forma, métodos e técnicas utilizados para desenvolver algoritmos que não apenas atendam às exigências técnicas mas também respeitem e enriqueçam a experiência humana, buscam proporcionar um equilíbrio entre inovação tecnológica e responsabilidade social, assegurando que os avanços em IA sejam acessíveis e benéficos para todos, independentemente de suas capacidades ou contexto social.

4

Metodologia

Neste capítulo, abordaremos a metodologia adotada para essa presente pesquisa exploratória, que visa aprofundar e compreender cenários que podem favorecer a reprodução de vieses de gênero no âmbito de aplicações de IA. A metodologia é estruturada em quatro componentes principais. Primeiramente, utilizaremos o mapeamento sistemático para identificar e analisar a literatura existente e as tendências prevalentes sobre o tema. Em seguida, realizaremos entrevistas com inquirição contextual com especialistas em IA, permitindo-nos explorar as nuances de suas experiências e percepções em relação aos desafios encontrados no desenvolvimento de algoritmos. Adicionalmente, serão conduzidos grupos de foco para facilitar discussões mais aprofundadas com alguns dos entrevistados, promovendo uma troca sobre os estímulos que alguns direcionadores baseados em Design Universal podem proporcionar. Por fim, a análise das co-ocorrências nos dados das entrevistas e grupos de foco será direcionada a destacar possíveis padrões e conexões relevantes, contribuindo para uma compreensão das dinâmicas subjacentes aos vieses observados.

4.1

Mapeamento Sistemático

O Mapeamento sistemático da literatura é um método para sistematizar o rastreamento e classificação de artigos científicos a partir de uma estratégia de busca conectada às perguntas de pesquisa especificadas.

4.1.1

Estratégia de busca

Mourão (2020) faz uma comparação entre diferentes estratégias de busca para realizar Revisões Sistemáticas da Literatura (SLRs-Systematic Literature Review) aplicadas ao campo de engenharia de software. O estudo aponta que o uso da estratégia híbrida combinando a busca de dados no repositório *Scopus* com o método paralelo de busca “*backward and forward snowballing*” utilizando o *Google Scholar*, tende a apresentar um equilíbrio adequado entre a qualidade dos resultados e a amplitude da revisão.

Com base nisso, a estratégia adotada é da busca híbrida, que envolve a aplicação de uma sequência de busca no repositório *Scopus*, a filtragem dos resultados usando os critérios de exclusão definidos e a realização de iterações simultâneas “*backward and forward snowballing*” nos estudos restantes (Mourão, 2017; Mourão, 2020). Ao seguir esse processo paralelo, os artigos obtidos pelo *backward snowballing* não foram sujeitos ao *forward snowballing*, e vice-versa.

Para definir a *string* de busca, que determina a sequência de rastreamento de artigos no repositório *Scopus*, foi utilizada a lógica de critérios PICO (Kitchenham, 2007), que se apresenta da seguinte forma: População (Population): Algorithms; Intervenção (Intervention): Gender Bias; Comparação (Comparison): N/A; Resultado (Outcome): N/A. Assim, a sequência de busca resultante foi “*algorithms*” AND “*gender bias*”. A *string* foi aplicada a títulos, resumos e palavras-chave no *Scopus* em outubro de 2023.

4.1.2
Seleção de artigos

O Critério de Inclusão (CI) e os Critérios de Exclusão (CEs) do estudo para filtrar as pesquisas são apresentados em Table 4.1.

Tabela 4.1: Critérios de Inclusão e exclusão	
Critério	Descrição
CI1	Descreve a etapa de modelagem algorítmica e relaciona questões sócio demográficas para a influência de vieses algorítmicos de gênero.
CI2	Aborda temas de contratação, divulgação de vagas de trabalho que refletem vieses algoritmos.
CE1	Não aprofunda no projeto do algoritmo ou avalia apenas os resultados.
CE2	Não foca em vieses algorítmicos de gênero especificamente.
CE3	Não aborda sistemas de recomendação e sim outras tecnologias como NLP, LLM e Computer vision
CE5	Artigos não publicados em conferências, jornais ou workshops <i>peer-reviewed</i>

4.1.3
Extração de Dados

Os dados extraídos de cada artigo compreendem metadados gerais das publicações (ano, autores, etc.), as respostas a cada questão de pesquisa e as vantagens e desvantagens relatadas para algumas das práticas. A Tabela 4.2 mostra o formulário de extração de dados. O formulário e todos os dados

extraídos também estão disponíveis em forma de planilha no documento disponível no repositório online ¹.

Tabela 4.2: Formulário de Extração de Dados

Informação	Descrição
Metadata	Título, Autores e Ano de publicação
Evidências de vieses de gênero gerados por algoritmos (QP1)	Exemplificação dos vieses
Evidências de práticas de explicabilidade ou similares para analisar para as decisões de projeto(QP1)	Indicação das técnicas e exemplificação da aplicação
Observações, Vantagens e desvantagens	impactos gerados pela aplicação da técnica de explicabilidade
Contextualizações sociodemográficas para descrever impactos nos resultados (QP2)	Evidências dos indicadores descritivos
Observações, Vantagens e desvantagens	valor agregado interpretado a partir dos indicadores sociodemográficos
Participação multidisciplinar nos projetos de modelagem algorítmica (QP2)	Descrição das equipes de projeto
Observações, Vantagens e desvantagens	valor agregado descrito a partir da experiência com equipes multidisciplinares

Foram analisados 14 artigos selecionados, a partir das seguintes perguntas:

- 1a. Há evidências de vieses de gênero gerados por algoritmos e exemplificação dos vieses, quais as implicações de vieses de gênero?
- 1b. O projeto do algoritmo é analisado e as decisões de projeto de modelagem do algoritmo são evidenciadas?
- 1c. Há Indicação das ferramentas e exemplificação da aplicação da técnica de explicabilidade, transparência, *responsible AI* ou *fair AI*?
- 1d. Qual técnica ou método utilizado para mitigação de vieses de gênero?
- 1g. Especifica problemas em sistemas de recomendação e envolve busca de emprego?
- 2a. Possui contextualizações sociodemográficas para descrever impactos e geração de vieses algorítmicos de gênero?
- 3a. Há a utilização de conceitos e técnicas de Design centrado no humano com indicação e descrição das técnicas?

Passamos à apresentação dos resultados encontrados (códigos e categorias) com trechos de respostas obtidas por meio de leitura diagonal e resumos

¹<https://shorturl.at/etzDG>

estratégicos. Utilizou-se de indicadores não frequenciais que permitem inferências de conhecimentos relativos às condições de produção das respostas Barbosa e Chen, 2019.

Foi realizada exploração de materiais, por documentos a posteriori, por dedução e codificação de registro (no caso temáticas por intensidade). Em seguida, prosseguiu-se a transformação das respostas que, por recorte, agregação e enumeração, permitiram a representação dos conteúdos ou suas expressões.

Os códigos foram elaborados por frases que capturassem simbolicamente a essência ou ecoassem um atributo para uma porção baseada na linguagem; esse código permite organizar e agrupar dados codificados similarmente em categorias, em decorrência dos padrões compartilhados. Esse padrão teve um rótulo para combinar os termos identificados numa mesma resposta e esse rol foi comparado pergunta a pergunta de cada artigo. O processo se desenvolveu em três etapas: definição, exemplificação de trechos e exemplificação de regras (alto, intermediário, baixo).

Posteriormente, foi feita a categorização cujos critérios foram léxicos baseados nas perguntas. Houve a exclusão mútua de perguntas não respondidas, ou que apresentaram dúvidas na resposta, indicando nova pergunta para geração de resposta, foi feita a análise da homogeneidade e objetividade da questão.

A Unidade de registro confere um significado e corresponde a um segmento de conteúdo considerado como unidade-base, cujo objetivo é categorizar unidades semânticas por tema. Esse tema foi identificado por núcleos de sentido que compõem a comunicação cuja presença ou frequência de aparição significa algo em relação ao objetivo do estudo. Então: exploração do material, codificação, recorte por unidade de registro e temática, enumeração por presença e frequência.

Assim, de todos os artigos, a saturação acima de 55% foi considerada relevante, sendo 19 artigos avaliados em 11 perguntas cada. Todas as perguntas foram analisadas em grupos de núcleos de sentido e homogeneidade da questão, os grupos de perguntas não respondidas foram excluídas, sem se agregarem a outros códigos.

No total, 209, grupos de respostas foram avaliados (1a,1b,1c,1d,1e,1f,1g, 2a, 2b, 2c, 3a), e em cada grupo, foi considerada uma saturação de até 100% de respostas viáveis. A categorização está na Tabela 4.1.3.

Tabela 4.3: Categorização

Codificações (Temáticas)	Unidades de Análise (Artigos Encontrados)	Percentual Temático
Evidências e vieses de gênero	Propagação de vieses (17)	98
	Retroalimentação de informações enviesadas (15)	89
	Predição de resultados tendenciosos (16)	92
	Predições enviesadas (17)	100
Implicação dos vieses	Impacto social desfavorável (14)	82
	Discriminação universal (9)	52
	Aprendizagem tendenciosa da IA (6)	35
	Inexistência de ferramentas de mitigação consensual (14)	85
Adaptação	Necessidade de filtragem e avaliação da IA (12)	87
	Dificuldade de mitigação de danos (8)	45
	Estratégias para adequação (12)	86

O texto abaixo tem por objetivo expor as relações entre os conteúdos dos artigos e as características dos autores eleitos para este estudo, qualitativamente, de forma dedutiva, a partir de questões organizadas previamente e do formulário de extração de dados (documento disponível no repositório on-line²). Para a determinação das categorias, considerou-se: (1) regras claras de inclusão e exclusão nas categorias, (2) categorias são mutuamente excludentes, (3) homogeneidade entre as categorias. Posteriormente, a partir dessas categorias, foram formadas unidades de análise para condensar os dados extraídos.

Com base na saturação das respostas analisadas, em que se buscou identificar as tendências e padrões dos dados extraídos, buscou-se observar o campo de forma mais ampla, ainda com a compreensão de que uma análise da totalidade dos dados seria impossível, em decorrência da extensão do campo. A partir dessa extração de dados, foram identificadas três principais categorias:

1. Evidências de vieses de gênero:

Foi identificada, em todos os artigos (100%), a presença de vieses em várias áreas aonde a IA foi usada. Observou-se, assim, um tipo de efeito cascata: isso porque, uma vez identificada a presença de vieses (ainda que involuntários) nos dados que alimentam as IAs, as IAs, agora influenciadas por esses vieses, passam a se tornar um dos principais propagadores de vieses (98%). Uma vez que os sistemas reforçam esses vieses, foi possível também observar uma retroalimentação de informações tendenciosas (89%). Foi evidenciado que a influência na predição de resultados tendenciosos (92%) e retroalimentação de vieses (100%) corroboram com a perpetuação das predições estruturadas enviesadas.

2. Implicação dos vieses de gênero:

A manutenção de preconceitos relacionados ao gênero, apontam um impacto social expressivo e desfavorável (82%) com a evidenciação da

²<https://shorturl.at/etzdG>

discriminação em vários ambientes, enfaticamente as de busca de empregos (52%) o que favorece a manutenção de sistemas tendenciosos que reforçam os estereótipos presentes na aprendizagem da IA (35%). Apesar de já haver estudos que discutem vieses de gênero, não houve um consenso para aplicação de ferramentas de eliminação ou calibração adequadas ou conciliadas entre os autores, apenas (35%) apresentaram possíveis utilizações de métodos, porém a grande desigualdade e os processos de mitigação ainda são primários e não há ferramentas de eliminação ou calibração que possam ser amplamente utilizadas (85%).

3. Adaptação:

Não houve possibilidade de contextualizar os vieses por técnicas de correção dos grupos analisados. Os autores indicam a necessidade da organização da aprendizagem da IA com filtragem e avaliação frequentes (87%). Existem claros desvios de design e modelagem, com a retroalimentação dos dados enviesados, apenas 3 programas foram citados o WES, GF e NFCF (23%), mas não atingem a demanda ampliada ou conseguem mitigar preconceitos analisados através de IA devido ao quantitativo e velocidade de informações geradas pela IA (45%). Como estratégias de adequação para mitigação ou equivalência, a calibração, o aumento do vetor e otimização (54%) devem estar presentes, e a IA deve ter um acompanhamento mais responsável e transparente (32%); essas medidas são essenciais para uma adequação e ajuste para uma reparação que consiga uma equivalência sensata e correta dos preconceitos e melhorar o impacto social da discriminação gerada pelos vieses.

Quanto à contextualização sócio-demográficas, 44,4% dos artigos abordam o tema de maneira direta, formando associações e fornecendo exemplos de relações com a produção e perpetuação de vieses. Embora alguns dos artigos não se debrucem especificamente sobre essas contextualizações, ou busquem analisá-las a fundo, ao menos reconhecem o papel do contexto sócio-demográfico na produção e propagação de vieses.

4.2

Entrevistas semiestruturada com adição do método de Inquirição Contextual e formulário online

A opacidade nos processos de formação do aprendizado de máquina é uma constante argumentação para que os sistemas inteligentes tomem

caminhos não desejados. As entradas e saídas podem ser definidas, porém, o processamento do aprendizado vai percorrendo camadas e tomando decisões inacessíveis ao conhecimento humano (Verganti et al., 2020). O desafio é entender os componentes de projeto, os contextos humanos e sociais por trás das decisões para captar quão influentes são as deliberações humanas, como estas podem ser alteradas pelo contexto de formação profissional, pessoal e questões sociodemográficas de quem está atuando e entender se norteadores que figuram inclusão e diversidade podem apoiar no balizamento das tomadas de decisão durante a fase de projeto.

Para aprofundar conhecimento sobre o desenvolvimento de algoritmos, há a pretensão de visibilizar a jornada dos dados, desde a definição dos requisitos, e principalmente nos pontos que compõe o ato de inferência para atingir um problema específico. O método de investigação do processo de desenvolvimento consistiu em entrevistar por meio do método de inquirição contextual um grupo de desenvolvedores homens e outro de desenvolvedoras mulheres para entender todo o processo de decisões e escolha de ferramentas que eles empregam para fazer o projeto de modelagem algorítmica. O conteúdo colhido nas entrevistas foi transcrito, categorizado utilizando o método de análise de conteúdo e depois foi realizada a discussão das classificações dos códigos por meio de dois grupos de foco, conforme podemos visualizar no infográfico 1.2.

Na entrevista foi proposto que os desenvolvedores simulassem o processo de desenvolvimento de um algoritmo para atuar na divulgação de um anúncio de uma vaga de emprego na área de tecnologia em uma empresa brasileira de médio porte. O algoritmo deveria rastrear os candidatos no *Linkedin*, uma plataforma intitulada como comunidade profissional, sendo um espaço em que profissionais criam seus perfis com dados profissionais, como formação acadêmica e experiência profissional.

Com a intenção de entender os momentos mais sensíveis e mapear possíveis contextualizações e desdobramentos entre experiências e referências pessoais nas decisões de projeto, foi pedido aos entrevistados que narrassem e justificassem suas escolhas. Havia a intenção de contextualizar as decisões tomadas a partir do que foi visto e ouvido durante a entrevista. Também foram cruzados os dados sociodemográficos coletados tanto no formulário de anuência de participação quanto em uma breve introdução da entrevista, em que foi pedido para que os participantes descrevessem um breve histórico de suas experiências profissionais.

O método escolhido para embasar a estratégia da entrevista será a Inquirição Contextual. O método de inquirição contextual é um método de pesquisa que busca analisar fenômenos complexos dentro de seus contextos natu-

rais (Holtzblatt et al., 2004). A inquirição contextual reconhece a interdependência entre elementos do sistema e procura entender como esses elementos se influenciam mutuamente. O método é baseado na premissa de que a realidade é altamente situacional e dinâmica, valoriza a coleta de dados em ambientes naturais, considerando as nuances das relações entre as variáveis e as influências ambientais. Isso permite uma compreensão mais profunda e holística do fenômeno estudado.

Participantes - recrutamento

Os entrevistados foram selecionados através da rede de contatos profissionais da autora. A abordagem foi feita por canais fechados de comunicação. A partir do aceite, foi enviado um formulário para a coleta dos dados demográficos e experiência profissional específicos. O formulário também continha o termo de consentimento e esclarecimentos éticos sobre a participação e sua anuência. Nesse formulário constavam as seguintes seções e campos:

1. Dados

- Termo de consentimento
- Nome
- Email

2. Questões demográficas:

Gênero

- Feminino
- Masculino
- Não binário
- Prefiro não responder

Idade

- Menos de 18 anos
- 18 - 24 anos
- 25 - 34 anos
- 35 - 44 anos
- 45 - 54 anos
- 55 - 64 anos
- 65 anos ou mais

Com qual raça você se identifica?

- Negra
- Indígena
- Parda

- Amarela
- Branca
- Não sei
- Prefiro não responder

Região do país de moradia

- Norte
- Nordeste
- Centro-Oeste
- Sudeste
- Sul
- Fora do Brasil

3. Especificações Profissionais**Formação Profissional**

- Graduação em Ciência da Computação / Tecnologia da Informação
- Graduação em Engenharia
- Graduação em Matemática
- Graduação na área de humanas
- Outra área (especifique)
- Outros. . .

Qual tipo de Universidade você cursou?

- Federal
- Particular
- Particular com FIES ou alguma bolsa
- Outros. . .

Ramo da Atividade Profissional Atual

- Educação
- Indústria
- Governo
- Outros. . .

Direitos dos participantes Os perigos associados à pesquisa incluem desconfortos potenciais ou restrições ao expressar opiniões. Para superar esses obstáculos, foi assegurado o anonimato das respostas, o que impede a identificação dos participantes no material e nos resultados do estudo. Aos voluntários foi enfatizado que não haverá julgamentos ou avaliações em qualquer fase, sendo o foco principal da entrevista entender como os

desenvolvedores se organizam mentalmente para tomar decisões de questões técnicas e não técnicas em relação ao desenvolvimento de algoritmos.

Os participantes terão acesso aos resultados do estudo e serão informados sobre a importância e a validade tanto da pesquisa quanto da sua contribuição. Eles foram conscientizados sobre as medidas de confidencialidade e sobre o direito de se retirar do estudo a qualquer momento, sem enfrentar penalidades ou constrangimentos. A temática da pesquisa e os responsáveis por ela foram claramente comunicados. Todas essas informações constam no formulário de consentimento informado, entregue antes do início da primeira etapa do estudo.

Suspensão ou encerramento da entrevista

O participante foi informado do seu direito de abandonar o estudo durante todo o processo sem comprovação, garantindo a proteção das liberdades pessoais. Qualquer risco ou possibilidade de dano que tivesse sido descoberto durante o estudo e não especificado anteriormente nos Termos de Consentimento Livre e Esclarecido seria tratado com suspensão imediata do estudo. A Resolução CNS 466/12 descreve os riscos de uma pesquisa, como sendo a “possibilidade de danos à dimensão física, psíquica, moral, intelectual, social, cultural ou espiritual do ser humano, em qualquer pesquisa e dela decorrente”. (BRASIL, 2012).

Condução e ambiente da entrevista

As entrevistas foram realizadas em ambiente virtual, de maneira individual, em que o desenvolvedor(a) estava preferencialmente em seu ambiente comum de trabalho. Após a explanação do problema, foi solicitado que a tela do participante fosse compartilhada, e as decisões e escolhas de ferramentas fossem narradas.

Foram realizadas 10 entrevistas com especialistas do gênero masculino e 10 com especialistas do gênero feminino. Os dados coletados na narrativa registrada por meio de imagem e som, foram transcritos utilizando a plataforma *turboscribe* e as unidades de registro e contexto foram categorizadas utilizando a plataforma *Dovetail*.

As entrevistas tiveram duração média de 40 minutos, e foram realizadas através do *Microsoft Teams*, plataforma que permite a gravação dos vídeos e transcrição automática. Os entrevistados foram selecionados por apresentarem características bem diversas em relação aos seus dados sociodemográficos, com gêneros declarados diferentes, porém com experiências profissionais e técnicas parecidas.

Foram elaborados direcionadores para pautar toda a entrevista. Os direcionadores, sendo perguntas abrangentes, foram projetados na tela conforme o estudo de caso foi sendo detalhado pela entrevistadora. A entrevista foi evo-

luindo conforme os slides foram sendo apresentados.

Roteiro da Entrevista O roteiro da entrevista foi dividido em cinco partes, detalhadas a seguir:

1. Contextualização

Pergunta: Pode nos contar um pouco sobre seu papel na empresa e sua experiência com algoritmos? Logo de início, durante a contextualização, será pedido aos entrevistados que detalhem a experiência profissional que tiveram com tecnologia de inteligência artificial em geral e seu papel na atual empresa.

2. Cenário

Descrição: Uma empresa de tecnologia de porte médio contrata você através de uma empresa de divulgação de vagas para anunciar a seguinte oportunidade e características:

- Vaga de tecnologia para profissional pleno;
- Skills: desenvolvedor/a full Stack;
- Formação: engenharia de software, tecnologia da informação, sistemas e afins;
- Experiência: 5 anos de experiência;
- Regime de trabalho remoto.

Foi apresentado esse cenário fictício, para que os entrevistados pudessem tangibilizar a experiência de desenvolvimento com a definição de alguns detalhes sobre o anúncio da vaga, porém sem nenhum limitante técnico ou financeiro. Os entrevistados foram encorajados a questionar as especificações necessárias para a realização dessa tarefa, à medida que foram enumerando o “passo-a-passo” escolhido por eles.

3. Simulação

3A Descrição: Planejamento e Concepção

- Definição de Requisitos: Como você definiria os requisitos para o algoritmo?
- Dados Utilizados: Que tipo de dados você utilizaria para alimentar o algoritmo? Como esses dados deveriam ser coletados e processados?

3B Descrição: Desenvolvimento

- Escolha de Métodos: Quais métodos ou técnicas de machine learning podem ser considerados para o algoritmo? Detalhe essa escolha?

- Desafios Técnicos: Quais os principais desafios técnicos deverão ser enfrentados durante o desenvolvimento do algoritmo?
- Integração e Implementação: Como o algoritmo foi integrado ao sistema de divulgação de vagas da empresa?

A terceira etapa se divide em dois momentos, planejamento e desenvolvimento. No primeiro momento, foi pedido aos entrevistados que narrassem como escolherão os requisitos e parâmetros para inserir no modelo algorítmico. Na sequência, os entrevistados foram estimulados a descrever a etapa de desenvolvimento.

4. Testes e Validação

Descrição:

- Testes Realizados: Quais tipos de testes poderiam ser aplicados para validar o algoritmo? (ex.: testes A/B, validação cruzada)
- Métricas de Sucesso: Quais métricas seriam usadas para avaliar o desempenho do algoritmo?
- Feedback e Iterações: Como avaliar feedbacks, interno ou externo, que podem levar a modificações no algoritmo?

5. Resultados e Impactos Descrição:

- Efetividade: Como você pode avaliar a efetividade do algoritmo na recomendação de candidatos com as skills desejadas?
- Impacto na Contratação: Como o algoritmo pode gerar algum impacto significativo no processo de contratação da empresa?

Nas duas últimas etapas, os entrevistados foram perguntados sobre as etapas de validação, testes e definição de métricas.

Análise e métricas de avaliação

O conteúdo das entrevistas foi transcrito e categorizado por meio de análise de conteúdo. Segundo Bardin (2009), a análise de conteúdo é um conjunto de técnicas abrangentes do campo das comunicações. Bardin ressalta que para haver análise de conteúdo em qualquer modo de comunicação, em que pese a existência de um emissor e um receptor, a mensagem deve ser decifrada por meios de escrita analítica. Para haver a descrição analítica são aplicados procedimentos sistemáticos direcionados para cobrir os objetivos de descrição do conteúdo das mensagens (Bardin, 2009). Portanto, há a extração da informação que responde aos questionamentos sistemáticos contidos no objetivo da pesquisa, expressado, por exemplo, por meio de entrevistas.

A sistemática da interpretação da mensagem se dá em três fases: pré-análise, exploração do material e tratamento dos resultados, por meio de inferência e interpretação. Na fase de pré-análise acontece a categorização com a formação de um esquema preciso para a exploração da pesquisa, em vias de viabilizar códigos para convergir com as hipóteses e objetivos iniciais (Bardin, 2009).

Após a pré-análise, as categorias qualitativas serão analisadas de maneira a buscar o reconhecimento de padrões e questões que evidenciam os modelos mentais formados para a resolução do problema proposto.

Segundo Bardin (2009), a inferência em análise de conteúdo se destina a delimitação dos conhecimentos relativos inerentes às condições da produção da mensagem. Desta forma, estas inferências – operações lógicas propostas baseadas em proposições aceitas como verdades – procuram responder a dois tipos de problemas já contextualizados na temática da pesquisa:

1. Quais foram as causas do enunciado (mensagem capturada) – foi realizada uma investigação das causas contextuais para as decisões de projeto do entrevistado(a), sobrepostas as evidências pessoais e questões socio-demográficas do participante.
2. Quais efeitos da mensagem – o fluxo das decisões tomadas foi categorizado para deixar os modelos mentais menos abstratos.

Após realizadas todas as entrevistas, os áudios foram transcritos utilizando a plataforma *turboscribe* e as unidades de registro e contexto foram categorizadas utilizando a plataforma *Dovetail*, aplicando a técnica de Análise das co-ocorrências, que representa a extração de relações entre os elementos de mensagem contidos em fragmentos de textos, onde se destacam simultaneamente dois ou mais unidades de registro na mesma unidade de contexto (Bardin, 2009). Em seguida, foi realizada a codificação para análise dos resultados, e os trechos de textos retirados das transcrições foram categorizados com palavras-chave que estão compreendidas nas temáticas da pesquisa: experiências pessoais, aspectos de UX, princípios de IA e estruturação técnica para projetos de algoritmos. A análise seguiu cronologicamente o roteiro da entrevista, marcado pela categoria de estrutura técnica e as relações com as outras categorias, conforme fossem classificadas, foram guiando a discussão dos resultados.

No contexto da entrevista, foi introduzida a ideia de um checklist para apoiar os especialistas a estruturarem os requisitos de seus projetos com uma visão expandida para necessidades específicas de usuários e tratativa de temas sensíveis, com possibilidade de geração de vieses de gênero, sociais e raciais. O

checklist foi aprofundado e discutido em dois grupos de foco, conforme descrito adiante.

4.3

Grupo de foco

Segundo Ressel et al. (2008), o método de grupo de foco reúne um grupo de participantes que recebem estímulos da pesquisadora(o) para promover debates com foco em tópicos diretivos. O ambiente deve ser ameno, ideal que não seja composto por participantes de um mesmo ciclo social para que as opiniões possam ser expressas livremente, sem receio de retaliação (Borges, 2005). Nessa dinâmica, a pesquisadora pode ouvir vários pontos de vista em uma mesma ocasião e a partir da exposição de um mesmo conteúdo. Desta forma, a intenção foi a geração de insights a partir da troca de informações, sentimentos e experiências diversas acerca do tema exposto (Kind, 2004), diminuindo o grau de parcialidade da interpretação da pesquisadora. Os participantes de um grupo focal devem compartilhar características relacionadas ao tema estudado, garantindo a homogeneidade do grupo em aspectos que influenciem significativamente a percepção do tópico. Barbour e Kitzinger (1999) sugerem a seleção de indivíduos familiarizados e com conhecimento profundo sobre o assunto em discussão, e os fatores que impactam os dados (Kitzinger e Barbour, 1999).

Participantes, recrutamento, condução e material

Foram realizados dois grupos de foco com seis e sete participantes cada. Ambos os grupos foram compostos por pessoas entrevistadas na primeira etapa e que tinham disponibilidade para participar da dinâmica online. O grupo de pessoas que participou da entrevista e do grupo de foco foi selecionado a partir da rede de contatos profissionais da autora que foi sendo desenvolvida ao longo dos anos. Foram selecionados profissionais com experiência em pesquisa aplicada em tecnologias como realidade aumentada, virtual e inteligência artificial, e profissionais da indústria atuantes na área de *Analytics and Artificial Intelligence*.

Em relação à composição de gênero dos dois grupos de foco, um foi exclusivamente composto por mulheres desenvolvedoras e outro por homens desenvolvedores. A sessão foi estruturada em torno da discussão de seis indicadores, todos ligados a um checklist desenvolvido com base em conceitos de desenvolvimento de produtos inclusivos. O principal objetivo dessa dinâmica foi investigar se o checklist torna o processo decisório dos desenvolvedores mais co-participativo, além de aumentar a confiança e o conforto desses profissionais ao tomarem decisões em projetos de impacto social direto.

Durante a sessão, cada indicador foi apresentado em voz alta para todos os participantes. Após a leitura de cada item, os participantes tiveram dois minutos para refletir e responder a cinco questões reflexivas baseadas em suas interpretações pessoais:

- O que você entende sobre esse item?
- Você acha realizável no dia a dia?
- Como você aplicaria este item?
- Quais meios buscaria para implementá-lo?
- Esse item te fez pensar em algo que não costuma considerar durante o desenvolvimento de algoritmos?

Essas perguntas e os indicadores estavam dispostas em um painel digital colaborativo na plataforma *Miro* e foram desenhados para estimular uma reflexão profunda sobre a aplicabilidade prática dos indicadores e sobre a inclusão de novas perspectivas no processo de desenvolvimento. Após o tempo inicial de reflexão e escrita sobre cada item, os participantes tiveram em torno de 10 minutos para discutir os pontos mais relevantes que lhes vieram. O grupo de foco também serviu como uma plataforma para que os desenvolvedores expressassem suas ideias e preocupações, promovendo um ambiente de aprendizado mútuo e de cocriação.

Os seis indicadores compilados e propostos por Kat Holmes em *Digital Suffragists* (Tessier, 2021), base da discussão do grupo de foco, são os seguintes:

1. Considere a diversidade sob a perspectiva das interações humanas e considere a maneira como as pessoas mudam ao longo do tempo;
2. Identifique vieses relacionados à capacidade, gênero, raça e/ou acessos e preveja possíveis interações incorretas que possam acontecer;
3. Crie uma diversidade de caminhos para a participação em sua experiência;
4. Assuma que o projeto terá incertezas, necessidade de iterações e adaptações e trabalhe com a utilização de habilidades complementares;
5. Desenvolva uma alfabetização básica em acessibilidade e se aprofunde nos critérios de acessibilidade específicos relevantes para suas soluções;
6. Adote uma definição mais flexível de projeto. Abra os processos e convide pessoas com habilidades relevantes a contribuir com seu projeto, mesmo que não sejam programadores (as).

A pesquisadora Kat Holmes do *MIT* levanta questões que permeiam o desenvolvimento de produtos que podem acentuar vieses de gêneros. Segundo Holmes (2021), alguns métodos ágeis de desenvolvimento de produto, como o *Minimum Viable Product (MVP)*, podem simplificar o entendimento do usuário, dificultando o desenho profundo de particularidades de gênero. Ela cita o exemplo do desenvolvimento do aplicativo *Apple Health* que mede vários parâmetros de saúde para avaliar o condicionamento físico do usuário. Questões como pressão, temperatura, características sanguíneas, etc. A primeira versão do aplicativo não considerava que pessoas que menstruam possuem uma variabilidade de certas condições medidas, que não se enquadram nos padrões que o sistema tomava como base e acabava acusando irregularidades nas medições não apropriadas para esses usuários específicos.

Análise da discussão

Direitos dos participantes Os perigos associados à pesquisa incluem desconfortos potenciais ou restrições ao expressar opiniões. Para superar esses obstáculos, foi assegurado o anonimato das respostas, o que impede a identificação dos participantes no material e nos resultados do estudo. Aos voluntários foi enfatizado que não haverá julgamentos ou avaliações em qualquer fase, sendo o foco principal do grupo de foco investigar se um checklist fundamentado em conceitos de desenvolvimento de produtos inclusivos pode tornar o processo decisório dos desenvolvedores/desenvolvedoras mais co-participativo, aumentando sua confiança e conforto para tomadas de decisão em projetos com impacto social direto.

Os participantes terão acesso aos resultados do estudo e foram informados sobre a importância e a validade tanto da pesquisa quanto da sua contribuição. Eles foram conscientizados sobre as medidas de confidencialidade e sobre o direito de se retirar do estudo a qualquer momento, sem enfrentar penalidades ou constrangimentos. A temática da pesquisa e os responsáveis por ela foram claramente comunicados. Todas essas informações constavam no formulário de consentimento informado, entregue antes do início da primeira etapa do estudo.

Suspensão ou encerramento da entrevista

O participante foi informado do seu direito de abandonar o estudo durante todo o processo sem comprovação, garantindo a proteção das liberdades pessoais. Qualquer risco ou possibilidade de dano que tivesse sido descoberto durante o estudo e não especificado anteriormente nos Termos de Consentimento Livre e Esclarecido seria tratado com suspensão imediata do estudo. A Resolução CNS 466/12 descreve os riscos de uma pesquisa, como sendo a “possibilidade de danos à dimensão física, psíquica, moral, intelectual, social,

cultural ou espiritual do ser humano, em qualquer pesquisa e dela decorrente”.
(BRASIL, 2012).

5

Análise de Resultados

Neste capítulo, será apresentada a análise detalhada dos resultados obtidos a partir de 20 entrevistas, conduzidas com profissionais da área de tecnologia, em que realizaram uma simulação de desenvolvimento de algoritmo com o tema proposto. As entrevistas foram realizadas com dez homens e dez mulheres, autodeclarados. Essas entrevistas foram codificadas em cinco categorias distintas, que delimitam a compreensão do modelo mental demonstrado pelos entrevistados. O primeiro passo para a criação de um modelo mental é a coleta de dados, e em seguida, aplica-se uma abordagem *bottom-up* para organizar comportamentos ou pensamentos semelhantes verticalizados (Young, 2008). Conforme a figura 5.1, as quatro categorias e subcategorias que formam os pilares do modelo mental estão divididas entre **Experiências Pessoais**, **Estrutura Técnica**, **Aspectos de UX** e **Princípios de Inteligência Artificial**. Depois da estruturação desses pilares, os produtos ou funcionalidades que derivam da discussão são alinhados abaixo dos pilares principais, alinhando-se às temáticas (Young, 2008), nesse caso trazemos os direcionadores que irão figurar o **Guia de Diretrizes para Algoritmos Sensíveis** que foram discutidos e analisados em dois grupos de foco, realizados em um grupo de sete mulheres e outro de seis homens, e, a análise dos resultados será apresentada no capítulo 6.

5.1

Experiências Pessoais

As **Experiências Pessoais** dos entrevistados oferecem um delineamento importantes sobre o impacto pessoal e profissional na escolha das tecnologias e metodologias discutidas. A descrição de suas experiências servem como base para o entendimento da condução prática, a medida que suas escolhas são narradas. Os códigos e suas respectivas descrições estão na tabela 5.1.

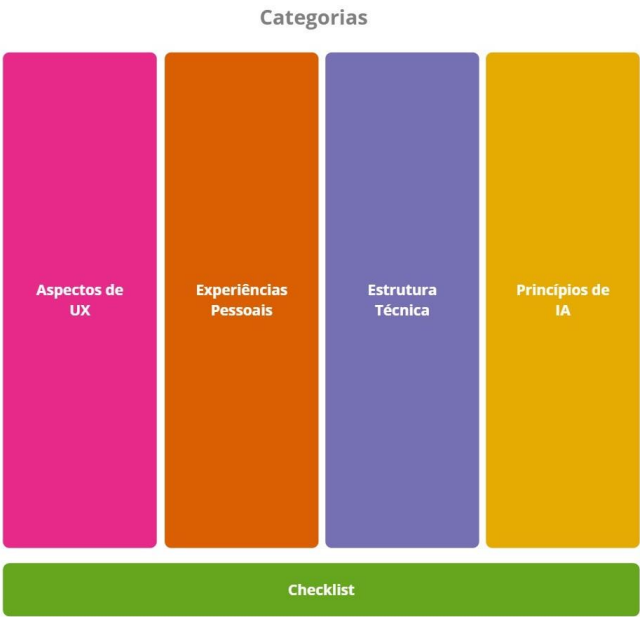


Figura 5.1: Categorias: estrutura de mapa mental

Tabela 5.1: Códigos para Experiências Pessoais

Código	Descrição
Idade	Indicação de alguma delimitação de tempo, em relação a acessos ou restrições
Gênero	Indicação de influência em decisões técnicas e/ou impacto na carreira relacionado ao gênero
Raça	Indicação de influência em decisões técnicas e/ou impacto na carreira relacionado à raça
Formação acadêmica	Descrição da formação e/ou algum impacto nas decisões técnicas em decorrência da experiência de formação
Região do país de moradia	Indicação de influência em decisões técnicas e/ou impacto na carreira relacionado a região de moradia

5.2
Estrutura Técnica

Para examinar as estratégias de resolução de problemas utilizadas pelos desenvolvedores em diferentes estágios do desenvolvimento algorítmico, identificar métodos adotados, métricas, definição de requisitos, entre outros, a categoria denominada **Estrutura Técnica** traz códigos para fazer a marcação desses indicativos. A roteirização da simulação de desenvolvimento proposta foi baseada na definição que traz a fase de modelagem e inferência se desdobrando em três etapas distintas: elencar os parâmetros que compõem esses modelos (1-Planejamento), identificar os modelos viáveis (2-Desenvolvimento) e avaliar o desempenho de cada modelo, recorrendo a métricas (3-Testes e Validação) e a um processo justo de comparação (Resultados e Impactos) (Kalinowski et al.,

2023). Desta forma, marcando um começo, meio e fim, os códigos que formam essa categoria de Estrutura Técnica estão classificados conforme a tabela 5.2.

Tabela 5.2: Códigos para estrutura Técnica. Fonte: desenvolvida pela autora.

Código	Descrição
1 - Planejamento	momento em que os participantes olham para o todo, entendem o desafio e capturam as regras de negócio. Indicação de aprofundamento e contextualização.
1a - Definição de Requisitos	O enunciado da simulação traz alguns requisitos bem alto nível. Esse código se refere a indicadores de questionamento de requisitos, contextualização e mostram o início da estratégia adotada pelos desenvolvedores.
1b - Dados Utilizados	Indicadores de demonstração técnica de como farão aquisição de dados que respondam os requisitos levantados.
2 - Desenvolvimento	Fase em que as questões práticas como escolha de métodos são aventados e escolhidos. Demarcadores que refletem a experiência dos desenvolvedores com algoritmos e preocupações em relação ao impacto social das suas escolhas técnicas, caso as tenham.
2a - Método	Indicação de escolha dos métodos, indicadores de viabilidade técnica e de discussão do alcance da tecnologia demonstrando o ponto de vista dos desenvolvedores em relação à viabilidade técnica.
2b - Desafios Técnicos	Indicação de alguma delimitação de tempo, em relação a acessos ou restrições.
2c - Integração e Implementação	Momento em que os desenvolvedores pensam na conexão das suas soluções com as plataformas digitais já usuais para o tipo de anúncio de vagas
3 - Testes e Validação	Fase de colocar em perspectiva suas aplicações e criar meios para testar as funcionalidades avaliando a capacidade, veracidade e adequação de suas soluções
3a - Testes	Métodos para testar suas aplicações, por meio de comparações com a intenção de validar a técnica aplicada.
3b - Métricas	Técnicas para marcar e evidenciar os resultados. Momento em que os desenvolvedores devem se colocar no lugar dos contratantes da solução e nas questões mais intrínsecas da temática para gerar marcadores que validem suas soluções.
3c- Feedback	Indicação que os desenvolvedores elaboram meios de capturar a opinião e evidências humanas e não humanas para avaliarem se a solução está contemplando as regras de negócio e atingindo as expectativas de resultado do cliente.
4 - Resultados e Impactos	Fase final em que os desenvolvedores olham para o todo e pensam diretamente no ponto de contato com os usuários, tanto quem consome, quanto quem contrata. Indicação de avaliação nos impactos individuais, ou seja, na vida de cada usuário, e também nos impactos coletivos, em sociedade.
4a - Efetividade	Demonstração que há a preocupação no entendimento se a aplicação de fato é útil e efetiva. Indicação de questionamento da tecnologia sob a perspectiva prática da atividade específica de recrutamento profissional de pessoas.
4b - Impacto na Contratação	Fechamento da simulação, com questionamentos e reflexões sobre o impacto em determinados grupos de pessoas, com entendimento de vantagens e desvantagens do ponto de vista social.

A partir da marcação desses pontos técnicos estruturais, utilizamos mais duas categorias para explorar os critérios e considerações éticas e sociais que influenciam as decisões dos desenvolvedores, visando entender como esses

fatores são integrados na prática de desenvolvimento.

5.3

Princípios de Inteligência Artificial

Na última década, praticamente toda a organização que trata de temas político-tecnológicos formulou ou endossou um conjunto de princípios para a Inteligência Artificial. Surgiram muitos documentos para balizar diretrizes para o desenvolvimento e consumo de aplicações com Inteligência Artificial de maneira ética, socialmente benéfica, salvaguardadas sobre uma governança de IA que engloba contextos de direitos humanos sobre o aspecto individual e coletivo.

Para analisar como os desenvolvedores processam os princípios éticos e socialmente sensíveis gerados pela automatização dos sistemas, utilizamos os oito temas mais citados em um mapeamento de guias para Princípios de IA e categorizados para essa análise como **Princípios de IA**. Esse mapa, desenvolvido em 2020, traz um mapeamento sistemático e visual dos documentos com abordagens éticas e questões de direitos humanos para IA (Fjeld et al., 2020). Os oito temas percebidos como padrões na totalidade de documentos compilados, funcionam nessa pesquisa como códigos para marcar as declarações dos desenvolvedores que demonstram sensibilidade a essas questões e seus impactos diretos aos usuários. Na tabela 5.3 temos a descrição dos códigos utilizados.

Tabela 5.3: Códigos para Princípios de IA.

Fonte: Tabela elaborada com conteúdo retirado de Fjeld et al (2020)

Código	Descrição
Equidade	Diante das preocupações globais sobre o viés em IA, os princípios de Justiça exigem que os sistemas de IA sejam projetados e utilizados de maneira a maximizarem a equidade
Impacto Social	Diante das preocupações globais sobre o viés em IA, os princípios de Não Discriminação exigem que os sistemas de IA sejam projetados e utilizados de maneira a promoverem a inclusão.
Privacidade e segurança	Os temas de Privacidade e Segurança em sistemas de IA enfatizam o respeito pela privacidade individual, incluindo o uso responsável de dados e o controle que as pessoas têm sobre seus próprios dados e decisões derivadas. Além disso, destacam a necessidade de os sistemas serem seguros e funcionarem conforme o planejado, protegendo-os contra acessos não autorizados.
Responsabilidade	Este tema aborda princípios relacionados à importância de mecanismos para garantir que a responsabilidade pelos impactos dos sistemas de IA seja distribuída de maneira adequada e que soluções apropriadas sejam fornecidas.
Explicabilidade	Os princípios deste tema estabelecem que os sistemas de IA sejam projetados e implementados para permitir supervisão, incluindo a tradução de suas operações para resultados compreensíveis e o fornecimento de informações sobre onde, quando e como estão sendo usados.
Ética Profissional	Esse princípio destaca a importância do profissionalismo e integridade dos envolvidos no desenvolvimento de sistemas de IA, enfatizando a necessidade de consultarem os <i>stakeholders</i> relevantes e planejarem os impactos de longo prazo desses sistemas.
Controle humano da tecnologia	Os princípios deste tema exigem que decisões importantes permaneçam sob revisão humana.
Promoção dos direitos Humanos	Esse princípio afirma que os objetivos e métodos de implementação da IA devem estar alinhados com valores fundamentais e promover o bem-estar da humanidade.

5.4

Aspectos de UX

Apresentaremos na sequência os **Aspectos de UX**, que se referem à Experiência do Usuário. Embora UX seja comumente aplicado para a avaliação de como os produtos são percebidos e utilizados pelos usuários, focando em elementos como usabilidade, acessibilidade e satisfação, os aspectos escolhidos para essa categorização estão focados em perceber momentos de dentro para fora, ou seja, quando os desenvolvedores têm ações e pensamentos direcionados para fazer escolhas que afetam diretamente os usuários finais. São considerados, portanto, cenários diversos para encaixar necessidades de diferentes personas, busca de estímulos e autonomia para o usuário, entre outros aspectos. As descrições dos códigos são baseadas nos conceitos de Marc Hassenzahl (2008) que elabora questões sobre a perspectiva da experiência do usuário sob a impressão da qualidade dos produtos (Hassenzahl, 2008), da qualidade das interações e do alcance dos objetivos diretos ou mais hedônicos durante as

interações. Wei Xu (2019) propõe um framework que direciona questões para promover uma boa experiência ao usuário, com tecnologia compreensível, explicável, útil e com boa usabilidade.

Desta forma, para essa avaliação de resultados, fizemos uma conexão entre esses aspectos de UX e os princípios de IA, funcionando como um “De Para” com um pensamento simplista e direto: da perspectiva do desenvolvimento das experiências, para garantir que os já referidos Princípios de IA (tabela 5.4) sejam experienciados, é necessário projetar a interação com o usuário final através da implementação de determinados aspectos de UX, detalhados na tabela 5.4. Essa relação entre os Princípios de IA e Aspectos de UX, irá auxiliar o entendimento da codificação das entrevistas e pode ser observada na imagem 5.2.

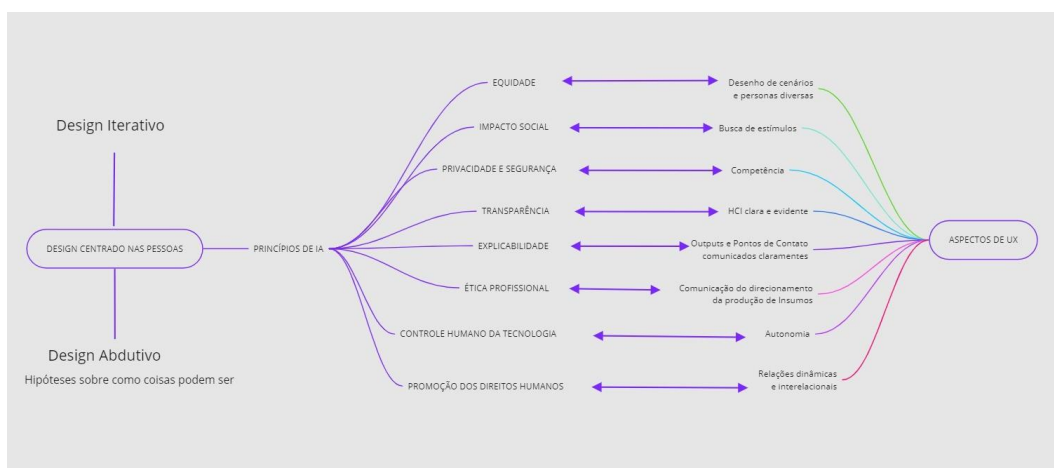


Figura 5.2: Esquema desenvolvido pela autora: relação entre Princípios de IA e Aspectos de UX.

Tabela 5.4: Códigos para Aspectos de UX.

Fonte: Tabela elaborada com base em Hassenzahl(2008)

Código	Descrição
Desenho de cenários e personas diversas	Indicação de que há preocupação em desenvolver soluções que atendam diferentes necessidades dos usuários, para isso há o exercício de projetar diferentes personas que serão os usuários finais de suas aplicações.
Busca de estímulos	Indicação de que há o planejamento de tornar sua aplicação mais palatável e com estímulos para que mais pessoas possam utilizar, tornando sua solução mais inclusiva.
Competência	O aspecto de competência no âmbito de UX é derivado da necessidade humana de sentir que tem as habilidades suficientes para controlar o uso da aplicação, para direcionar suas atividades (<i>self-oriented</i>). A conexão com Privacidade e Segurança (Princípio de IA) deriva do uso responsável dos dados e consciência e controle dos usuários sobre suas decisões.
HCI clara e evidente	Indicativo de que os desenvolvedores pensam em desenvolver soluções que tornem a relação humano-computador (UI, UX) clara de maneira que as responsabilidades e impactos sejam evidentes, comunicados e distribuídos de maneira adequada, e, em caso de prejuízos, estes sejam comunicados previamente.
Outputs e Pontos de Contato comunicados claramente	Indicador de que os desenvolvedores projetam as soluções pensando em garantir que todo o dado ou informação que derivem dessas interações sejam utilizados de maneira responsável e que estas ações e resoluções sejam comunicadas, rastreáveis e compreensíveis.
Comunicação do direcionamento da produção de insumos	Indica que há a preocupação em ter a anuência dos usuários em relação à utilização dos dados criados a partir do uso dos sistemas.
Autonomia	Dar ao usuário, tanto intermediário, quanto final, a possibilidade de direcionar ajustes e realizar revisões. Projetar soluções que permitam ao usuário manter o controle e escolha de uso.
Relações dinâmicas e inter-relacionais	Indicativo de que houve um pensamento formulando diferentes maneiras de organização em sociedade, de maneira a garantir que o uso do sistema esteja alinhado com valores que promovem os direitos humanos.

5.5

Guia de Diretrizes para Algoritmos Sensíveis

A última categoria compreende os indicadores para um desenvolvimento inclusivo, os seis indicadores compilados a partir de um checklist proposto por Kat Holmes em *Digital Suffragists* (Tessier, 2021), para o delineamento de ações e pensamentos inclusivos em desenvolvimento de produtos. No entendimento do modelo mental dos desenvolvedores, esses indicadores são como produtos que abarcam as consequências de outros pensamentos inclusivos, indicados pelos códigos de outras categorias. A discussão sobre o uso desses indicadores como direcionadores está no capítulo 6, na análise de resultados dos grupos de foco. Os códigos derivados do checklist estão na tabela 5.5.

Tabela 5.5: Códigos do Guia de Diretrizes para Algoritmos Sensíveis

Código	Descrição
Diversidade nas interações humanas	Considere a diversidade sob a perspectiva das interações humanas
Modificações ao longo do tempo	Considere a maneira como as pessoas mudam ao longo do tempo
Vieses de gênero, raça e agravantes de limitações	Identifique vieses relacionados à capacidade, gênero, raça e/ou acessos e preveja possíveis interações incorretas que possam acontecer;
Caminhos para a mesma experiência	Crie uma diversidade de caminhos para a participação em sua experiência;
Metodologia Ágil	Assuma que o projeto terá incertezas, necessidade de iterações e adaptações e trabalhe com a utilização de habilidades complementares;
Habilidades complementares	Adote uma definição mais flexível de projeto. Abra os processos e convide pessoas com habilidades relevantes a contribuir com seu projeto, mesmo que não sejam programadores (as).
Alfabetização em acessibilidade	Desenvolva uma alfabetização básica em acessibilidade e se aprofunde nos critérios de acessibilidade específicos relevantes para suas soluções;
Autonomia	Dar ao usuário, tanto intermediário, quanto final, a possibilidade de direcionar ajustes e realizar revisões. Projetar soluções que permitam ao usuário manter o controle e escolha de uso.
Relações dinâmicas e inter-relacionais	Indicativo de que houve um pensamento formulando diferentes maneiras de organização em sociedade, de maneira a garantir que o uso do sistema esteja alinhado com valores que promovem os direitos humanos.

Os códigos foram colocados como marcadores em trechos transcritos das 20 entrevistas. Alguns trechos com apenas um código e outros com conjuntos de códigos, que podem gerar análises mais abrangentes sobre a amplitude do envolvimento e decisões dos desenvolvedores ou demonstram um pensamento isolado. Para analisar a estratégia de desenvolvimento que deriva em modelos mentais, será primeiramente apresentado um compilado de informações sobre as experiências pessoais para que o perfil dos entrevistados fique mais claro. Após, serão apresentadas as principais relações semânticas entre os grupos de códigos, seguindo a sequência de perguntas roteirizadas. Desta forma, os códigos de **Estrutura Técnica** norteiam a passagem de tempo das entrevistas e junto a eles, destacamos algumas interpretações e percepções com a marcação também, dos códigos das categorias de **Aspectos de UX e Princípios de IA**.

5.6

Análise das entrevistas: Experiências Pessoais

Os entrevistados foram abordados por canais de mensagem como *WhatsApp*¹ ou *Microsoft Teams*² e depois das primeiras entrevistas, foi compreendido que o envio do formulário com questões pessoais de formação, idade,

¹<https://www.whatsapp.com/>

²<https://www.microsoft.com/pt-br/microsoft-teams/group-chat-software>

gênero e uma breve explicação da pesquisa, deveria ser enviado após a realização das entrevistas, para não enviesar as respostas. Algumas pessoas pareciam entender que deveriam direcionar as suas respostas uma vez que na descrição da pesquisa havia descrito o recorte de gênero do estudo. Desta forma, ao iniciar a entrevista, o termo de consentimento era lido e feita a anuência para gravar a imagem e áudio da entrevista.

Os entrevistados foram selecionados através do método bola de neve, a grande maioria foi contato profissional da pesquisadora. Houve uma maior dificuldade para completar o número de desenvolvedoras, então contamos com indicações para atingir o número necessário. Sobre os dados demográficos como região de moradia, idade e raça que se identificam e algumas questões sobre a formação dos participantes, chegamos no panorama da figura 5.3.

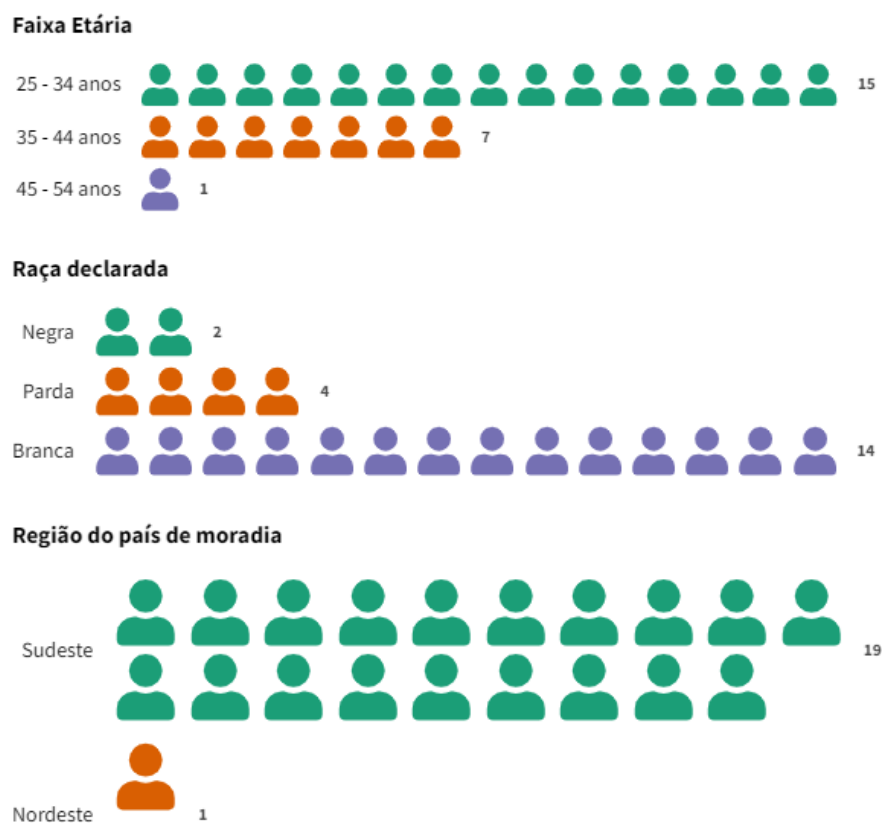


Figura 5.3: Dados demográficos dos especialistas. Fonte: Figura elaborada pela autora

Os dados refletem um panorama bem comum para pessoas que trabalham com tecnologia, a grande maioria de 70% está concentrada na faixa etária entre 25 e 34 anos, três pessoas se declararam pardas e apenas uma se identifica com a raça negra. Do total, 66,7% tiveram formação da graduação em universidades

federais e 72,2% trabalham na indústria. Na figura 5.4 podemos observar o detalhamento da especificação profissional dos participantes e do vínculo das respectivas universidades.

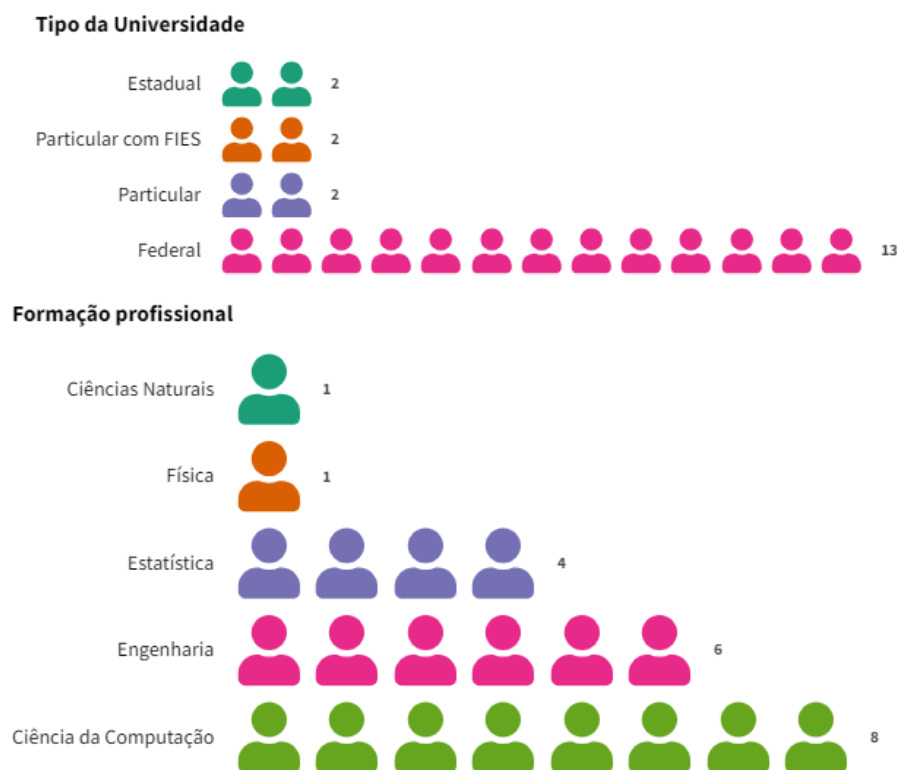


Figura 5.4: Formação acadêmica dos especialistas

Conforme observamos, nenhum participante teve sua graduação em áreas de humanas ou outras áreas que não fossem de exatas. Porém, durante a entrevista, muitos falaram sobre a alta recorrência de pessoas que fazem transição de carreira, porém nesse grupo, apenas duas pessoas tiveram suas graduações fora do plano de carreira mais comum para desenvolvedores, sendo em física e geofísica.

Ao interpelar os candidatos para uma possível seleção à participação nas entrevistas, o principal requisito era que tivessem experiência em modelagem algorítmica, sem necessidade de ter experiência em algum método específico, mas que já tivessem trabalhado profissional ou academicamente desenvolvendo algoritmos. Diante desse requisito, alguns desenvolvedores não tiveram confiança em participar da entrevista, em geral, as mulheres tiveram mais dificuldade em se declararem confiantes com a experiência que tiveram, mesmo sendo apontadas por outros profissionais como pessoas com experiência com modelagem algorítmica, algumas desistiram de participar.

Na figura 5.5 observamos a distribuição dos participantes conforme o nível de experiência declarado e a quantificação em anos de suas experiências profissionais.

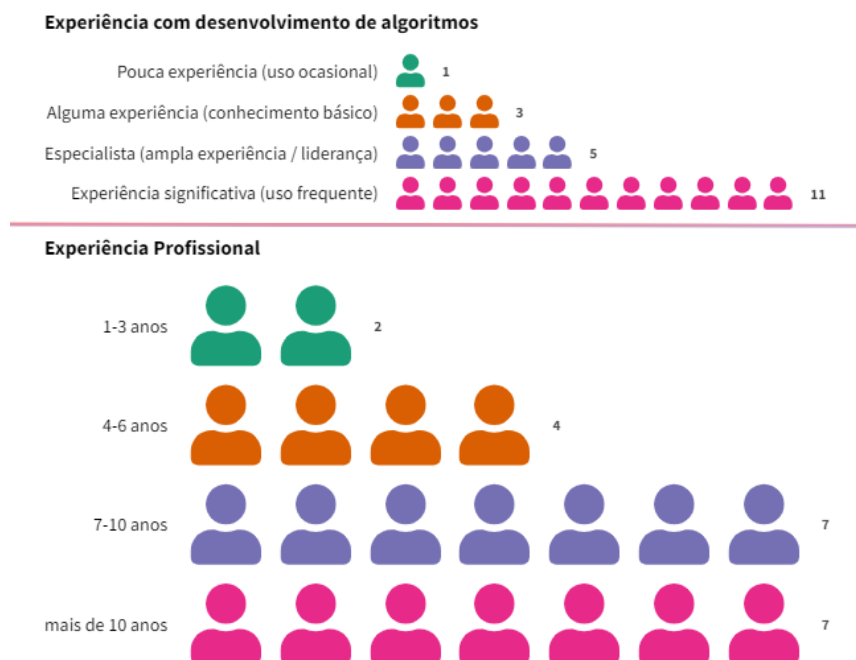


Figura 5.5: Experiência profissional dos especialistas

No formulário havia um campo aberto para que compartilhassem informações adicionais sobre suas formações, caso desejassem, e uma das participantes escreveu sobre ter feito cursos para entrar na área de dados em um local que ofereciam programas específicos para mulheres e pessoas negras.

Ao realizar as entrevistas, o primeiro direcionamento era para que compartilhassem sobre suas experiências profissionais e acadêmicas com desenvolvimento de algoritmos e sobre seus papéis nas empresas atuais, caso estivessem empregados. No geral, todos os desenvolvedores acabaram direcionando as suas escolhas técnicas durante a simulação feita na sequência, para o campo que mais tinham domínio, porém dois valem ser destacados aqui por refletirem claramente nas suas justificativas as suas experiências durante a formação. Rebeca, faixa etária entre 35-44 anos, declarou: “Eu sou formada na Unesp em estatística. Lá a gente tinha pouco acesso ao computador. Eu não tinha computador na época também até por questão financeira e tudo mais. Lá a gente teve uma matéria, uma disciplina, que era de Pascal. Foi a primeira vez que eu comecei a ter contato com qualquer coisa de programação.”

Ao se deparar na simulação com o requisito para o anúncio da vaga de cinco anos de experiência, Joana questiona: “...quando eu saí da faculdade,

eu não tinha computador, aí eu fui cair no mercado de trabalho assim, eu tive a sorte, que estatística ainda tava sendo bem contratada.... Mas hoje em dia, uma pessoa que se formou lá na Unesp, como eu, que não tinha computador, tá muito atrás daquela galera que tem a oportunidade de ter um computador, tem tudo em casa. E aí, quando eu tô criando uma vaga... seja esses cinco anos de experiência, não necessariamente pode ser cinco anos de experiência no profissional, ela pode ter um conhecimento que ela estudando foi lá fazendo durante a faculdade. Então, eu tô sempre selecionando a galera que é privilegiada.”

Já num campo mais técnico, Roberto, faixa etária entre 35-44 anos, trouxe para a simulação a solução técnica de testes rebatida a partir de sua ex- experiência acadêmica. Durante o desenvolvimento de seu mestrado e doutorado em computação, João trabalhou com algoritmos normativos para inteligência artificial, que segundo ele eram desenvolvidas normas para direcionar os algoritmos a atuarem ou não, de maneira automática, na resolução de determinados problemas. Ou seja, os algoritmos deveriam seguir normas, se não as respeitassem eram punidos. então, para a solução de calibragem do desenvolvimento do algoritmo simulado na entrevista, João sugere utilizar um currículo guia com derivação de normas a serem seguidas para que o algoritmo encontre currículos similares e faça a seleção.

Um dos entrevistados, após o término da entrevista, quando recebeu o formulário e leu sobre o tema da pesquisa, ele procurou a pesquisadora e disse estar envergonhado por não trazer temas de diversidade e inclusão durante a simulação, e afirmou que esses são temas muito caros a ele porque, inclusive, se formou com o fomento de políticas públicas de financiamento.

Uma das entrevistadas, algumas semanas depois da entrevista, buscou a pesquisadora para dizer que vinha refletindo sobre a importância da discussão de iniquidade de gêneros em ambientes profissionais de tecnologia ser discutido entre as equipes e não vir somente pela liderança. Muitos dos entrevistados, ficaram entusiasmados no geral por poderem pensar e discutir os caminhos para a resolução do desafio proposto.

5.7

Análise das entrevistas: Estrutura Técnica, Aspectos de UX e Princípios de Inteligência Artificial

Após a contextualização de suas experiências pessoais, os desenvolvedores foram apresentados ao desafio que consistia em narrar suas escolhas para o desenvolvimento de um algoritmo que deveria selecionar automaticamente alguns usuários de uma plataforma digital para visualizarem um anúncio de

vaga de emprego em uma empresa de tecnologia de médio porte com os seguintes requisitos:

- Vaga de tecnologia para profissional pleno;
- Skills: desenvolvedor/a full stack;
- Formação: engenharia de software, tecnologia da informação, sistemas e afins;
- Experiência: 5 anos;
- Regime de trabalho remoto.

A análise seguirá o roteiro, conforme refletido na categoria de estrutura técnica, onde as unidades de análise funcionam como âncoras para, a partir da correlação com outras unidades de análise, formarem unidades de contexto (trechos) codificados. O esquemático da análise pode ser visualizado na figura 5.6.

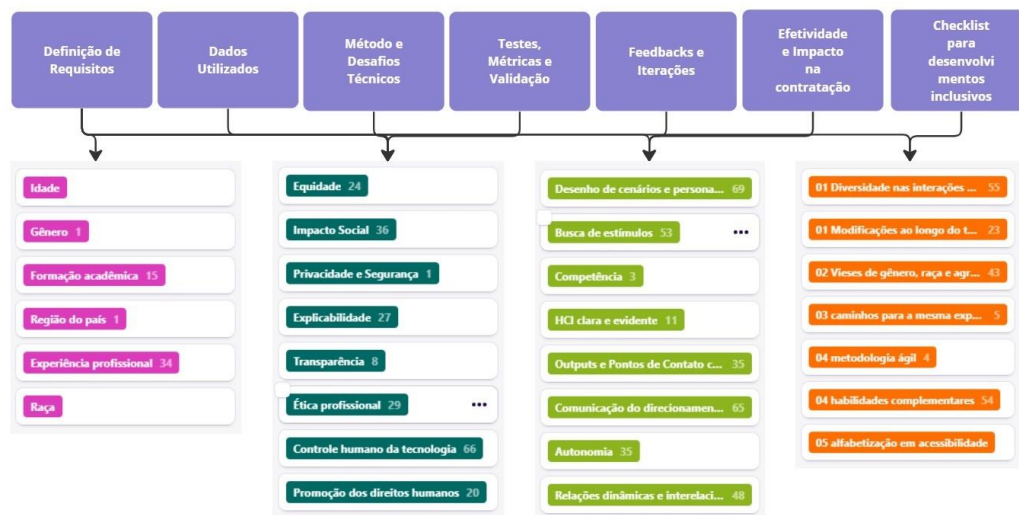


Figura 5.6: Esquemático da análise de resultados

5.7.0.1

Planejamento: Definição de Requisitos

Ao primeiro contato com o desafio todos os entrevistados apresentam dúvidas e necessitam de mais informações, já que a definição dos requisitos é apresentada de maneira bem alto nível, para que os questionamentos surjam, e gerem discussões com a explanação de seus pensamentos.

Ao fazerem o exercício de aprofundamento para ter os requisitos melhor estruturados, do grupo de homens, sete já introduzem falas que circundam

questões da experiência do usuário, como busca de estímulos, desenho de cenários e personas diversas, comunicação dos outputs e pontos de contato, trazem lições aprendidas de suas experiências profissionais, introduzem a aplicação de técnicas de explicabilidade e sobre controle humano das tecnologias e destacam a importância de definirem bem as regras de negócio com pessoas especialistas da área específica de RH (habilidades complementares).

Do grupo das mulheres, três apresentam nesse momento falas sobre aspectos de UX como cenários diversos, relações dinâmicas e inter-relacionais e comunicação da utilização dos insumos.

Nos trechos seguintes, Roberto, Renato, Francisco (nomes fictícios) falam sobre ouvir quem está demandando a solução, através de uma escuta ativa, para ter completo esclarecimento sobre as regras de negócio. Nesse trecho foram marcadas, além das âncoras (seção), as categorias: habilidades complementares, busca de estímulos, relações dinâmicas e inter-relacionais, comunicação do direcionamento da produção de insumos.

“você vai ter que entender o problema dela (pessoa), verificar se aquilo de fato é um problema que dá pra você gerar (...) um software ou algo que vai solucionar o problema dessa pessoa. A partir desse ponto, você vai começar a direcionar essa pessoa de forma que ela consiga mostrar pra você o que ela precisa em forma de funcionalidades. Então, assim, nós temos, por exemplo, a possibilidade de sentar pra pessoa e falar assim: Olha o que você imagina nesse seu eh sistema? E aí você deixa a pessoa falar ela explicar o que ela quer, o que ela gostaria. (...) essa, pra mim tem se demonstrado ser a (solução) mais eficaz ao nível de mercado e aí, por exemplo, com relação à parte de dados, né? Feito esse levantamento dos requisitos, você tem a possibilidade de criar ali uma primeira versão, né? Rápida, pra mostrar pro seu cliente validar (...) se tá fazendo sentido, né? É a mensagem que ele te passou, então você pode usar softwares, que você não precisa programar, apenas fazer o desenho ali de telas pra mostrar (...) um fluxo do sistema.” (Roberto)

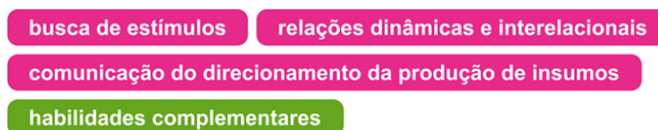


Figura 5.7: Trecho 1. Figura elaborada pela autora.

“...tentar fazer algum levantamento com o cliente (...) ou sobre o conhecimento da empresa, também, de histórico, de quais seriam as melhores formas de fazer essa segregação, né?” (Renato)

habilidades complementares

Figura 5.8: Trecho 2. Figura elaborada pela autora.

“Eu ia simplesmente perguntar pro contratante (...) o que eles estão tentando fazer, que eles sentem que não tem o conhecimento e (...) se eles pudessem me dar um exemplo de alguma coisa que eles precisam fazer, algum caso em específico, e que eles não têm o conhecimento ou a capacidade pra fazer pra que eu pudesse utilizar isso nas minhas entrevistas”. (Francisco)

busca de estímulos

comunicação do direcionamento da produção de insumos

habilidades complementares

diversidade nas interações humanas

Figura 5.9: Trecho 3. Figura elaborada pela autora.

Já mais sobre o âmbito prático de extração dos requisitos, nos dois trechos seguintes Afonso e Rebeca trazem questões que gostariam de debater com o demandante (fictício) para clarearem as regras de negócio pensando em suas próprias experiências e, projetando experiências diversas.

“Porque, por exemplo, pleno com cinco anos de formação, talvez, teria uma dificuldade de encontrar tá? Talvez seria interessante entender porque são cinco anos de formação pra um pleno, né? Então, por exemplo, no meu caso, eu tenho cinco anos e sou sênior, outras pessoas também que estão nessa faixa já são sêniores, né?”

desenho de cenários e personas diversas

diversidade nas interações humanas

modificações ao longo do tempo

Figura 5.10: Trecho 4. Figura elaborada pela autora.

“Então, eu acho que só a formação ela não leva a trazer, às vezes, a melhor pessoa, né? Pode ser que a pessoa tenha tido uma experiência muito maior do que a formação em si. Então o histórico de carreira mesmo, (...) olhar todo o histórico profissional dessa pessoa.” (Rebeca)



Figura 5.11: Trecho 5. Figura elaborada pela autora.

“Eu acho que vem da minha formação, do pessoal que eu trabalhei, eu normalmente, desenho aonde eu quero chegar. Pra mim, encontrar a pessoa, é a característica final, então eu faria o reverso e delimitaria quais são as variáveis que Faz sentido eu ter em relação a cada uma dessas variáveis que a gente tem como delimitador.” (Caio)



Figura 5.12: Trecho 6. Figura elaborada pela autora.

Por fim, Vinícius é o primeiro a inserir questões de gênero na fase de definição de requisitos:

“Informações sobre gênero, não colocaria nenhum marcador específico pra gênero. Pode ser pessoa de qualquer gênero, porque é independente.” (Vinícius)



Figura 5.13: Trecho 7. Figura elaborada pela autora.

5.7.0.2

Planejamento: Dados utilizados

Essa fase, sobre a definição dos dados a serem utilizados, é a segunda mais densa de relatos (39m12s), a primeira é a fase final com a discussão de resultados e impactos (56min27s). Nessa seção estão reunidos trechos mais relevantes selecionados a partir dos marcadores âncoras da categoria de estrutura técnica: planejamento, definição de requisitos e dados utilizados. Houve uma distribuição bastante homogênea de marcadores entres ambas as falas de homens e mulheres, em que trazem preocupações do âmbito técnico para extração de dados, mas colocando questões de cuidado na consideração e tradução ética e justa dados em informações. Seguem-se alguns pontos-chave extraídos das entrevistas:

– Equilíbrio entre Machine Learning e decisões humanas:

Alice expressa preocupação com a exclusividade dos algoritmos em processos de seleção, sugerindo que, apesar das eficiências proporcionadas por *machine learning*, é preciso manter abordagens tradicionais para garantir que aspectos humanos não sejam ignorados, como a sensibilidade às nuances que os algoritmos podem perder.

“Então eu acho importante que tenha essa questão do machine learning, né? Porque é algo que é para facilitar(...) só que eu não acho que deveria ser exclusivo a isso, porque no fim das contas a gente está lidando com pessoas, né? Então, eu vejo que às vezes, dependendo da forma que esse algoritmo vá ser construído, isso pode ser mais excludente, né? Então, eu acho que eu uniria duas formas, uma mais tradicional, que é literalmente a gente coloca os requisitos, recebe a candidatura da pessoa e avalia, ou a pessoa dos processos do RH, né? E a outra parte seria com machine learning, ver se tem match das palavras ou match dos requisitos, né? Então, assim, eu optaria por seguir essas duas frentes de maneira paralela, porque eu... na construção do algoritmo, a gente tem um processo que pode ser, não falho, mas pode ser realmente limitante. Às vezes eu tenho uma pessoa muito boa e na hora de escrever, por exemplo, ela não colocou a palavra chave X. Só que ela foi uma pessoa que vai se adequar a esse perfil, né? Então, se eu sou uma pessoa no RH, que eu estou procurando uma pessoa, que ela não vai ser apenas uma caixinha... eu também tenho que abrir os meus olhos para isso.”
(Alice)



Figura 5.14: Trecho 8. Figura elaborada pela autora.

– Critérios de Seleção e Experiência:

Denise e Luciana discutem como os critérios de seleção podem ser ajustados para incluir uma variedade de fatores relevantes, não apenas a experiência tradicional ou os títulos acadêmicos. Elas apontam para a necessidade de adaptar os algoritmos para refletir uma gama mais ampla de competências e experiências.

“eu trabalho com linguagem, e (olho) num espaço setorial similar, que sejam temas afins com o desenvolvedor fullstack, porque vai ter muita coisa que é análogo. Então eu também colocaria como demanda desse algoritmo ter um match na formação desses temas que me interessem e eu colocaria como critério. Eu tô entendendo que experiência não é um critério flexível. Também, né Eu colocaria como corte. Assim, se tiver menos de cinco anos de experiência, é um sim ou não, assim, uma árvore de decisão e não selecionar se isso for mais flexível.” (Denise)



Figura 5.15: Trecho 9. Figura elaborada pela autora.

“Necessariamente eu vou falar que mais anos de experiência é melhor? Não necessariamente a gente vai fazer isso, então olhar para quem a gente já contratou pra essa vaga, porque geralmente (...) quando a gente faz um algoritmo supervisionado, você coloca alguns casos de sucesso como referência. Quem você contratou, quais eram as informações que batiam? Então, por exemplo, a formação, a gente vai ter várias formações diferentes.” (Luciana)



Figura 5.16: Trecho 10. Figura elaborada pela autora.

“Só que aí você começa a entrar em alguns problemas também nisso. Por exemplo, é lá atrás na minha carreira era muito comum os filtros de empresas pegarem assim, formou na USP, na Unicamp, na UFRJ e o resto cortava,(...) claramente você tá deixando um público de fora e quem disse que esse público de fora não é tão bom quanto o outro? Então, o ponto é, eu tenho muita gente aqui boa que tem essa formação de escola de elite aqui dentro, mas o quanto outras pessoas de outras escolas não encaram. Então, mesmo que eu faça ali um algoritmo caixa preta, eu só coloque a saída, coloque os meus casos de sucesso e a saída sem eu conseguir ver qual os pesos das variáveis, eu tenho que fazer depois uma análise dos perfis para ver que perfis estão sendo selecionados para entender quais variáveis estão pesando.” (Luciana)



Figura 5.17: Trecho 11. Figura elaborada pela autora.

Mariana e Roberto, já trazem visões sobre o peso da formação acadêmica e a natureza da universidade. Mariana coloca que mesmo ela tendo formação em universidade federal, e entender o valor. Roberto considera importante não fechar nesse espectro e Roberto considera necessário a exigência de ter graduação concluída.

“(...) porque não necessariamente um profissional que vem de uma universidade particular, ele não é competente pra atender aquela vaga, né? É claro que quando a gente vem de faculdades federais ou faculdades com nome, a gente tende a entender que profissionais dessa mesma bolha vão atender o que você procura. Mas eu, Camila, eu não fecharia os olhos pra entrevistar pessoas

de outras faculdades, de outras formações, eu prezo muitas coisas na formação. Possivelmente, vai ter muita gente de formação, eu acredito tá, de formação em federal, porque a área de tecnologia, ela é bem forte nas universidades federais, agora tá aumentando nas particulares,(...) por exemplo, até 2021 não tinha curso de graduação em estatística em faculdade particular, só nas federais, mas eu acho que eu pesaria mais no quesito formação e, claro, na hora de selecionar os currículos, vamos supor que venham 200, não dá pra entrevistar duzentas pessoas! Vamos fazer um ‘mix’ aqui entre gênero e idade, e entrevistar pessoas com idade mais jovem, mas também com idades mais avançadas, porque tem pessoas que fazem migração de carreira e tá tudo bem.” (Mariana)

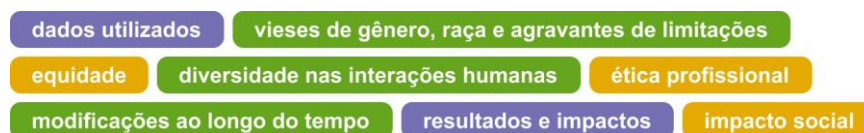


Figura 5.18: Trecho 12. Figura elaborada pela autora.

“Não seria um preconceito, mas assim, (é necessário) a pessoa ter uma graduação é porque assim, as pessoas que trabalham na área de tecnologia e não possuem graduação, elas têm muita dificuldade pra demonstrar o que elas sabem de fato. Ou que pelo menos elas têm a base necessária pra aprender novas tecnologias. Então, isso eu vejo como uma dificuldade.” (Roberto)



Figura 5.19: Trecho 13. Figura elaborada pela autora.

Nesse contexto, de dar maiores pesos a determinados requisitos, Luciana ressalta a relevância de manter o controle humano da tecnologia e poder de revisão, conforme o seguinte trecho:

“Por exemplo esse exemplo que deu de qual a universidade que se formou porque essa é uma informação que eu tenho aí no topo do funil que a pessoa colocou no cadastro do currículo. Eu quero que isso esteja na minha informação, pois está diferenciando um pouco,

talvez faça sentido, mas está diferenciando muito mais que todo o resto. Toda vez que uma variável tem peso no modelo, eu tenho que olhar para essa variável com crítica dela.” (Luciana)



Figura 5.20: Trecho 14. Figura elaborada pela autora.

– Recomendação e Similaridade:

Francisco e João falam sobre o uso de sistemas de recomendação para identificar candidatos ou produtos similares, o que pode ser útil tanto em recrutamento quanto em vendas, destacando como os algoritmos podem ser adaptados para classificar padrões e preferências, sem deixar de atentar para as questões éticas do uso de dados:

“seria uma estratégia de recomendação possível pra esse caso. (...) isso é muito utilizado pra produto, né? Qualquer coisa que você quer oferecer pras pessoas, você mapeia pessoas similares a ela, né? Em termos de tudo, dados demográficos, salário, não sei o quê você pode. Você pode acabar fazendo uma recomendação mais certa, né? Por exemplo, (...) se você gostou da bola de futevôlei, você pode gostar da rede de futevôlei também, né? E como eu vou saber quem joga futevôlei? Provavelmente quem tá nas cidades de praia e aí vai aí você com os outros dados, você pode alcançar mais facilmente pessoas similares, né? (...) Então, um cara que tem, dez anos de experiência, vai tá mais próximo de um cara com dez anos de experiência, um cara que é de Python, provavelmente vai tá mais próximo de um cara de Python, um cara que trabalhou na mesma empresa, sei lá, na Rede Globo, ele vai tá mais próximo de alguma maneira, porque naquela Feature nome das empresas (vai estar que) trabalhou na Globo. Então eu vou aproximar as pessoas numericamente de acordo com as suas características, né, com esses dados deles, né? Isso é uma maneira de buscar. (...) Eu

vou achar perfis mais próximos e, naturalmente, a idade vai ser uma coisa relevante, né? Então é, é bom também, antes de utilizar essa matriz de dados, você verificar se os dados são permitidos utilizar, né? E se faz sentido você colocar eles ali naquela nessa cesta, né? E quais dados você deve ocultar e utilizar e quais dados você não precisa nem mesmo ocultar, né? Por exemplo, identificar quem já é conectado com outra pessoa, né?” (Francisco)

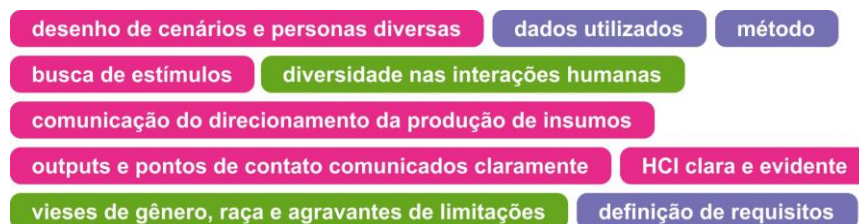


Figura 5.21: Trecho 15. Figura elaborada pela autora.

“É porque se você tem candidatos que não atendem os requisitos mínimos colocados, que são mais amplos, mas que não atendem eles, acho que não entraria. E aí, a partir disso, talvez clusterização, conseguir clusterizar esses grupos, essa população de indivíduos a partir das principais skills e aí conseguir trazer um panorama, um mapa mais definido pro cliente, né? Pra quem tem interesse em fazer essa contratação, por exemplo, você pode ter profissionais que tem uma ênfase numa determinada área, que é pertinente a esse escopo aberto da vaga e outros que tem uma ênfase maior em uma outra área que também se enquadraria nesse perfil mais amplo. Então, acho que fazer um processo de classificação pra entender onde tá essa população de indivíduos que faria sentido pra mim” (João)



Figura 5.22: Trecho 16. Figura elaborada pela autora.

– Alternativas mais amplas para extração de dados

:

Algumas pessoas também elucubraram sobre possibilidades mais abrangentes pra extrair dados para extrair padrões dos possíveis candidatos. Com intuito de dar mais chances para os candidatos se expressarem, ou simplesmente para buscar características auxiliares, porém refletindo situações que estão sob o domínio de conhecimento de cada um. Os seguintes excertos trazem esses recortes:

“... no Instagram e nas redes sociais, eu acho que estaria mais associado as páginas que essas pessoas seguem. Então, ah, se é uma pessoa que trabalha com TI, ela vai seguir páginas que são afins a essa área, né? Então vai seguir ali alguma página de memes de área de TI?” (Ana)

dados utilizados

definição de requisitos

planejamento

Figura 5.23: Trecho 17. Figura elaborada pela autora.

“Então, acho que isso e pensando mais um pouco como é algo um se esse cenário talvez tenha algum cultural, digamos assim, você pode seguir algumas instituições ali no LinkedIn. Então, talvez pegar instituições que estejam aderentes ao cultural, a empresa, um filtro inicial, digamos assim, mas que não é nada muito certo.” (Pedro)

caminhos para a mesma experiência

dados utilizados

busca de estímulos

Figura 5.24: Trecho 18. Figura elaborada pela autora.

“Mas no pleno, olhar é muito como a pessoa abordou aquela oportunidade, né? Tem pessoas que não conseguem aproveitar muito aquele campo de descrição, né, do que atuou naquela empresa. Isso acaba dificultando, mas eu olharia bastante aquele campo de descrição. Então, se a pessoa liderou, tá como pleno, mas lidera algum projeto ou foi responsável por desenvolver algum projeto dentro da empresa.” (Afonso)

– Gênero e Diversidade no Recrutamento:

dados utilizados

desenho de cenários e personas diversas

Figura 5.25: Trecho 19. Figura elaborada pela autora.

Alguns entrevistados mencionam questões sensíveis a serem abordadas como questões de gênero, indicando uma necessidade de estratégias de recrutamento que se atentem à inclusão e equidade.

“Mas se eu pudesse não enviar (o gênero), eu não enviaria. Mas, como em geral, as empresas depois pedem, então eu pegaria pra ter isso. Inicialmente, enviaria apenas a capacidade e o conhecimento técnico da pessoa. E depois eu enviaria essa informação porque em algum momento eu iria ter que enviar(...) A candidata ali se fosse o mais genérica possível em relação a gênero, etnia, religião ou coisas mais pessoais, assim. Então, porque independente de gênero, a gente tem que avaliar a capacidade da pessoa. Mas algumas empresas não necessariamente olham pra isso, tem muito preconceito. Então assim, se tivesse essa liberdade, eu ia dizer: ah, não vou pegar a foto, não vou pegar nome, nada disso.” (Pedro)

dados utilizados

vieses de gênero, raça e agravantes de limitações

relações dinâmicas e interrelacionais

desenho de cenários e personas diversas

Figura 5.26: Trecho 20. Figura elaborada pela autora.

“Tá o que era uma pessoa apta ou não, que exclusivamente atende aqueles requisitos e, se e se as pessoas que atendem aqueles requisitos da forma que eu especificar, de fato, são pessoas que tão aqui nessa bolinha, né? E a gente não tá olhando o resto, o todo. Então isso seria um desafio legal assim, de estudar E até que ponto vale usar informações pessoais também, né? Pra treinar esse modelo (...) eu não usaria informações como gênero, por exemplo. Mas se eu usar, pode ser que o modelo me retorne um resultado mais diversificado porque ele vai conhecer perfis, perfil de homens e de mulheres, né? Se eu não cito o gênero como uma feature, como uma variável de treinamento, ele simplesmente vai ignorar esse fato.” (Mariana)



Figura 5.27: Trecho 21. Figura elaborada pela autora.

5.7.0.3

Desenvolvimento: Método e desafios técnicos

– Método: customização e precisão de Algoritmos

Os desenvolvedores trouxeram algumas estratégias de desenvolvimento de algoritmos, métodos como classificação, árvore de decisão, regressão logística e segmentação e NLP (Neural Language Processing) foram citados, geralmente escolhidos, por já terem alguma experiência anterior. Foi solicitado para que a todos os entrevistados narrassem suas escolhas de maneira mais espontânea possível, e se tivessem que mudar suas escolhas, não teria nenhum problema, o mais importante era de fato passar o encadeamento das escolhas e decisões de maneira mais natural possível. Portanto, alguns dos desenvolvedores demonstraram alguma insegurança no primeiro momento, mas de maneira geral, iam construindo o pensamento com tranquilidade, lembrando de outras experiências para basearem suas escolhas naquele momento.

Bruno descreve o uso de um algoritmo de classificação, seguido por uma busca exaustiva de currículos para encontrar candidatos que correspondam a critérios específicos. Este processo considera treinar o algoritmo com um conjunto robusto de dados para garantir a confiabilidade dos resultados:

“Pode ser um algoritmo de classificação. E depois que eu tenha o meu algoritmo treinado, iria procurar com algoritmo de busca exaustiva de currículo pra procurar pessoas com aquela característica, baseado numa classificação que eu vou fazer baseado em um parâmetro que eu defini lá para as features pra esse meu classificador, seja o tempo de formação, tipo de formação, tempo que está empregado. Logicamente, que se for supervisionado eu já vou treinar ele com um conjunto de treinamento, validação e teste formatado que já saiba quais são as métricas do meu algoritmo e tudo mais. Eu já sei qual será a confiabilidade do resultado dele previamente, tá? Então eu já sei o que eu quero entregar e levar até uma um nível de confiança X. Mas vamos supor que o LinkedIn me pediu pra criar um modelo pra pesquisar, pra ajudar uma pessoa

de Rh. Ah! Vamos criar aqui um serviço pra facilitar a vida deles a encontrar pessoas com certa característica? Então eu, logicamente, teria que fazer um serviço dentro do LinkedIn, que desse oportunidade da pessoa colocar os dados que ele quer da vaga ali e esse algoritmo vai procurar match bom graças a seu treinamento. Eu tô batendo no ponto de que é um modelo treinado! Então o modelo já foi treinado pensando em todas as variações possíveis, tá? Então a pessoa vai preencher ali os seus dados, aquele modelo vai ali passar aquela busca exaustiva já no próprio LinkedIn. E aí, vai pra separar o que ele acha que da match (combina) com aquela característica, aquele requisito que a pessoa colocou e aí vai devolver um resultado pra essa pessoa.” (Bruno)

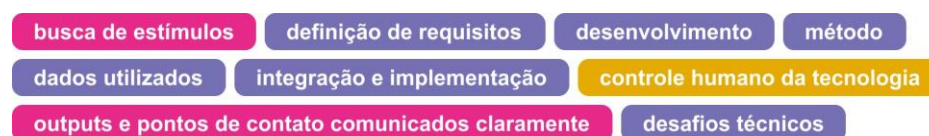


Figura 5.28: Trecho 22. Figura elaborada pela autora.

Mariana destaca a utilidade das árvores de decisão para classificar candidatos baseados em características específicas, como ano de formação ou tipo de graduação. Essa abordagem permite uma análise detalhada e uma filtragem eficaz de candidatos de acordo com as necessidades específicas da vaga:

“Ele (árvore de decisão) é muito bom pra esse tipo de classificação, porque a ideia é classificar o profissional, né? Do quão ele me atende? Tá, pensando naquela descrição de vaga, né? Então, a árvore de decisão ajuda porque cada raminho da árvore seria uma variável, né? Uma feature, e aí suponhamos que a feature seja ano de formação, aí eu quero saber se atende ou não, ou graduação que a pessoa fez será que atende diante das features que eu escolhi? Então, a árvore seria um bom método pra classificar esse profissional entre ele estar apto ao que você escolheu ou não.” (Mariana)

Luciana discute a prática de testar vários algoritmos, como regressão logística, árvores de decisão e redes neurais, para identificar o método mais eficaz em dadas circunstâncias. Isso sublinha a flexibilidade e adaptação na aplicação de técnicas de Machine Learning:

desenvolvimento

método

Figura 5.29: Trecho 23. Figura elaborada pela autora.

“Bom, aí daria para usar, sei lá, uma regressão logística, que é um método de classificação que costuma se adaptar bem a esse tipo de cenário, uma árvore de decisão ou uma rede neural, como hoje está cada vez mais fácil aplicar esses métodos, geralmente a gente faz vários e vê qual funciona melhor, mais do que escolher o método a priori. A gente faz um conjunto de métodos. A definição antes é, eu quero fazer previsão, eu quero fazer classificação? Aqui a classificação seria próximo passo ou não. Ou um score para ordenar quais são os melhores casos ali. E aí, por exemplo, uma regressão logística poderia me dar esse score. Aí, roda vários algoritmos e qual funciona melhor.” (Luciana)

controle humano da tecnologia

método

Figura 5.30: Trecho 24. Figura elaborada pela autora.

Joana menciona a segmentação para buscar perfis próximos e utiliza técnicas de *encoding* para extrair valor semântico de palavras, estratégia para tratar dados não estruturados em processos de recrutamento:

“Talvez, fazer uma segmentação pra buscar perfis próximos e talvez usar alguém que já esteja na empresa e que se aproxime do perfil, porque a gente não tem uma base de dados assim com label, né? De quem é um potencial desenvolvedor ou não. Então acho que talvez eu partiria pra essa parte da segmentação e aí pra filtrar, pra tratar esses dados, como a gente tá falando de dados não estruturados, talvez pra isso, utilizaria alguma técnica de *incoding* pra buscar ali o valor semântico dessas palavras próximas do que a gente tá procurando.” (Joana)

método

Figura 5.31: Trecho 25. Figura elaborada pela autora.

Davi discute a experiência do usuário (UX) na coleta de dados, destacando estratégias utilizadas para capturar informações relevantes conforme o ritmo de prazos curtos de atenção nas interações, facilitando a análise de similaridade entre candidatos:

“Os dados são esses que eu falei, pra você pegar esses dados, depende de como a sua aplicação vai interagir. Depende, Isso é mais uma questão de UX, mesmo tipo de uso da aplicação em si, da rede social ou do aplicativo, não importa o que for, pra poder pegar dados dos usuários daquela rede, pra digamos, ter esse cálculo de similaridade. Tô falando assim, pela experiência que eu tenho normalmente, o que os caras fazem hoje, (...) eles fazem você responder coisas rápidas, né? Porque hoje todo mundo tá incentivado a ver vídeo curto. Tudo curto, até os forms são curtos. Hoje, então, o direcionamento é que você entre o mais rápido possível na plataforma, com poucos dados e, durante o seu uso, você vai capturando outras informações que sejam importantes pra poder fazer esse tagueamento. (...) Ah, qual o melhor algoritmo? Existem várias formas. Eu posso, digamos, ranquear isso por palavras-chave, usar palavras, como formação, tempo de formação. Se formar é igual, eu consigo calcular isso ranquear, digamos, em tabelas, em bancos de dados e poder fazer uma similaridade.” (Davi)

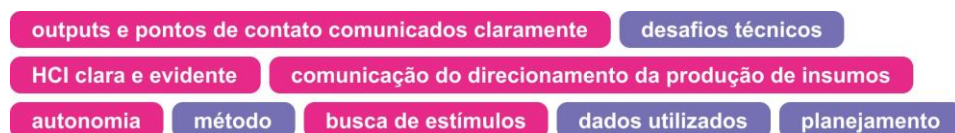


Figura 5.32: Trecho 26. Figura elaborada pela autora.

Renata explora o uso de engenharia de recursos (features) e NLP para extrair informações detalhadas de textos, como o tempo médio de experiência de trabalho e a utilização de expressões chave, para melhorar a precisão dos modelos de classificação:

“Você coloca o início e o fim, então poderia usar alguma em feature engineering pra pegar essas features relacionadas a tempo de trabalho naquela empresa, naquela organização pra poder fazer o tempo médio de experiência da pessoa. Então teria que fazer alguma coisa um pouco híbrido, então features engineering relacionadas a capturar esse tempo e algumas técnicas de NLP pra processamento de

texto. E aí você pode transformar esse texto em variáveis categóricas pra poder usar no modelo de classificação, de ranqueamento alguma coisa assim. Mas pra mostrar fulano é o potencial, então apresentar isso, pra capturar as palavras-chave, então construções também dá pra usar um engram pra poder capturar as construções de frases? Porque às vezes não é uma palavra, é uma expressão que é mais utilizada. Então também usaria isso pra processar o texto.” (Renata)



Figura 5.33: Trecho 27. Figura elaborada pela autora.

Caio fala sobre a progressão para algoritmos mais sofisticados, como redes neurais, quando características específicas apresentam alta probabilidade e os métodos mais simples não fornecem decisões claras:

“*Machine learning* ou uma árvore de decisão, alguma coisa nesse sentido e se eu continuasse ainda com uma dificuldade de tomar uma decisão em relação à probabilidade muito alta pra determinadas características eu iria pra um outro tipo de algo que já entraria no algoritmo mais profundo, como um algoritmo de rede neural.” (Caio)



Figura 5.34: Trecho 28. Figura elaborada pela autora.

– Desafios técnicos

Finalizando o segundo momento da estruturação técnica, conforme o roteiro seguido para a entrevista, os desenvolvedores trazem suas impressões sobre o que poderiam ser os maiores desafios técnicos.

1. Obtenção e Criação de Dados:

Ambos os entrevistados, Vinicius e Bruno, destacam a dificuldade em obter dados de qualidade e em quantidade suficiente para treinar os algoritmos. Isso inclui a necessidade de criar dados de forma artificial ou extrair informações de bancos de dados para garantir que as características essenciais estejam representadas nos dados de treinamento.

“Principais desafios técnicos no desenvolvimento do algoritmo seria a obtenção dessas informações (...) porque você precisa de um grande número de dados pra você criar um padrão. E aí você precisaria ter muitas pessoas com essa mesma característica, que a empresa tá buscando, né? Então, ou você cria esses dados, essas informações de forma artificial pra você treinar a rede ou você, então, busca essas informações em algum banco de dados, ou mesmo um banco de informação que você consiga integrar. Eu acho que esse é o maior desafio, porque a implementação mesmo não é, com o pessoal que tá já atuando na área, não é tão complexa assim. Se tratando de dados que vão virar dados, estruturados, né? São dados não estruturados, mas que viram dados estruturados depois.” (Vinicius)

método dados utilizados desafios técnicos

Figura 5.35: Trecho 29. Figura elaborada pela autora.

“Acho que o mais de desafiador, nem é a implementação do próprio algoritmo. Eu acho que seria a extração de todas essas features que eu vou usar pra alimentar ele. Ou você vai ter que ter um sistema de como eu já falei, de busca exaustiva por currículo pra definir quais são aquelas palavras-chave que vão ser utilizadas ou que seriam as features no caso. Ah, você tem que ser capaz de encontrar nelas, em qualquer um dos currículos, ou boa parte deles, pra você ter features suficiente pra todos os casos. Você sabe que a pessoa nem preenche corretamente os dados na hora que se cadastra nas redes sociais, nem todas as informações estão lá.” (Bruno)

desafios técnicos método dados utilizados integração e implementação
outputs e pontos de contato comunicados claramente busca de estímulos

Figura 5.36: Trecho 30. Figura elaborada pela autora.

2. Metodologia vs. Processamento de Dados:

Caio menciona a disparidade entre a variedade de metodologias disponíveis e a falta de robustez nos dados e na arquitetura tecnológica. Ele destaca que, embora existam várias formas de tomar decisões, a eficácia dessas decisões muitas vezes depende da qualidade e do processamento dos dados disponíveis.

“Então eu acho que eu até brinco isso com isso, com o pessoal que, às vezes a gente tem muita metodologia, tem muita maneira de se fazer, tem várias decisões, várias formas de tomar decisão, mas o que a gente não tem é processamento, ou não tem robustez de dados, ou não tem arquitetura tecnológica, que eu acho que a gente tá avançando e avançou muito. Mas a gente ainda tá avançando nesse sentido.” (Caio)

desafios técnicos

Figura 5.37: Trecho 31. Figura elaborada pela autora.

3. Base de Treinamento:

Mariana fala sobre construir uma base de treinamento sólida e representativa para ensinar os modelos de IA a identificar o que significa ser apto ou não para uma determinada vaga. Ela ressalta a complexidade de criar essa base, já que a qualidade dos dados de treinamento viabiliza a precisão do modelo.

“Como você vai treinar esse modelo? Como você vai ensinar a ele? O que é ser apto ou não? Isso aí é um desafio. Por quê? Porque eu posso construir essa base. Mas aí eu vou construir essa base com base no que sai da minha cabeça? Tipo, chegou uma pessoa lá na base, João, formado em tecnologia da informação, se formou em 2015, aí eu que vou dar resposta, eu preciso construir essa base pra treinar o meu modelo e automatizar isso. E aí eu acho que um dos maiores desafios seria ter essa base de treinamento” (Mariana)

método

dados utilizados

controle humano da tecnologia

Figura 5.38: Trecho 32. Figura elaborada pela autora.

Sobre os desafios de implementação, Luciana afirma a importância de auditar os passos durante a implementação de um algoritmo. Luciana destaca que é preciso verificar cada etapa do processo para garantir que, ao final, a classificação seja precisa.

“Acho que o desafio de implementação é fazer essa auditoria dos passos, porque eu vejo muitas vezes o negócio ir lá no final, chega lá no final a classificação não tá tão boa, mas eu vou ver... desde o

começo eu não tô conseguindo ler que aquela pessoa tem um skill de método ágil. Então, vou ter que, na implementação, garantir que cada parte da implementação está tendo o último output devido.”
(Luciana)

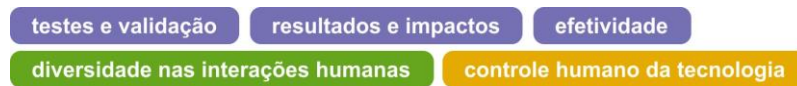


Figura 5.39: Trecho 33. Figura elaborada pela autora.

5.7.0.4

Testes, métricas e validação

Chegando na terceira fase da entrevista, os desenvolvedores foram questionados sobre técnicas para testar e validar os algoritmos. Diante do contexto da simulação proposto, os entrevistados deveriam entregar os algoritmos como produtos para seus contratantes, desta forma, é necessário testar para verificar se os algoritmos não estariam alucinando, errando, e se, de fato, estariam alcançando os resultados planejados. Portanto, a partir dos marcadores âncoras da categoria de estrutura técnica: testes, métricas, feedback e iterações, alguns trechos foram retirados para evidenciar as técnicas selecionadas. De uma maneira geral, os desenvolvedores trouxeram a preocupação em escutar especialistas ou outros colegas de equipe para discutirem estratégias e validarem seus métodos, demonstrado nos marcadores de habilidades complementares atribuídos a quase todos os trechos da temática. Também destaca-se a atenção geral a explicabilidade, autonomia e controle humano da tecnologia. A maioria desses trechos trazem esses referidos marcadores, em que os desenvolvedores discutem sobre a relevância de gerar descrições, ou algum tipo de rastreabilidade como mecanismo para ajustar possíveis resultados fora do planejado. Bem como, sempre manter o poder de decisão, direcionamento e correção com controle humano, e sobretudo, de especialistas na temática trabalhada.

– Testes e erros iniciais

Vinícius argumenta sobre testar modelos de ML internamente antes de aplicá-los em contextos reais. Embora ele reconheça que testes internos possam ser enviesados devido ao conhecimento prévio da equipe sobre os processos, essa familiaridade também facilita a identificação e correção de erros óbvios e graves.

“Então, dentro da própria equipe, dentro da própria empresa, você poderia criar um teste pra ver se realmente tá refletindo isso, como uma medida de segurança pra que não comece a disparar anúncios pra pessoas que não tem nada a ver. Com esse tipo de informação (...) você faria um teste com a própria equipe, seria um pouco enviesado esse teste. Porque você estaria fazendo com pessoas que já entendem como que funciona o fluxo. Entretanto, isso facilita pra correção de erros. Porque pelo fato dela entender como que o fluxo tá fluindo, como que são as saídas, você consegue identificar de uma forma mais fácil erros esdrúxulos, né? Problemas muito latentes! Assim que você fala é um erro muito crasso. Então volta e recomeça.” (Vinícius)

habilidades complementares

explicabilidade

Figura 5.40: Trecho 34. Figura elaborada pela autora.

– Explicabilidade e Transparência

Afonso enfatiza a necessidade de compreender como os modelos de ML chegaram às suas conclusões, o que ajuda a evitar a “caixa preta” e melhora a aceitação dos resultados por parte dos usuários e stakeholders. A explicabilidade é necessária para ajustes e melhorias contínuas.

“Se o modelo está indicando essas pessoas porque essas pessoas estão com score (nota) alto, aí tem técnicas de explicabilidade de ML que te ajudam a entender como que chegou na naquela resposta (...) pra explicar aquele resultado que o modelo entregou pra não ser uma caixinha preta. Isso é um problema muito antigo (...) que agora tá começando a melhorar com essas novas técnicas de explicabilidade, e a partir dessas técnicas você buscar entender porque chegou naquele resultado e fazer as devidas alterações”. (Afonso)

“Uma coisa que eu sempre cobro muito do meu time é que a análise não pode sair da analista de dados. A gente tem que saber mais o negócio para conseguir discutir. Então, não é pegar a task (descrição da demanda) e pegar pronta. Eu tenho que fazer uma segmentação.(...) ...não dá pra colocar essa base e eu vou te dar um notebook no Databricks que você roda e vai cuspir quem são



Figura 5.41: Trecho 35. Figura elaborada pela autora.

as pessoas dela. Você tem que discutir com a pessoa, por que quem vai entender mais de seleção? Eu espero, que a pessoa de RH entenda mais de seleção do que eu, por mais que, vamos supor, eu atendesse o people Analytics, que é uma área de expansão dentro de Analytics.” (Luciana)



Figura 5.42: Trecho 36. Figura elaborada pela autora.

– Métricas de sucesso com conhecimento do Negócio

Luciana, Mariana, Denise e Luciana valorizam o envolvimento de especialistas do domínio de negócios no processo de análise de dados. Isso assegura que as análises estejam alinhadas com as necessidades e objetivos reais do negócio e que as decisões baseadas em dados sejam pragmáticas e aplicáveis.

“A primeira versão desse modelo que normalmente não é utilizada, tá é só a primeira versão, o primeiro resultado. Eu chamo alguém que entenda do negócio pra dizer se faz sentido. (...) Por exemplo, no mundo dos acidentes. Faz sentido o modelo dizer que no turno da manhã tem mais acidentes do que no turno da tarde? Porque às vezes essa informação pra mim é super relevante no modelo, mas pro cara (especialista) vai dizer que não faz sentido, a equipe é a mesma, então eu envolvo bem sim a pessoa do negócio, mas assim após ter um primeiro resultado.” (Mariana)

“...para qualquer teste, melhor dos mundos é quando a gente consegue isolar um grupo teste-controle, né? Se você separar no teste A-B, seja rodar algumas vagas com e algumas vagas sem esse

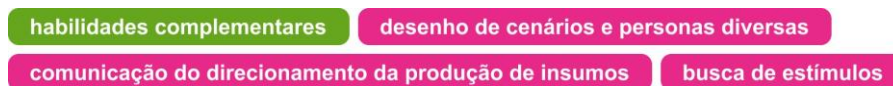


Figura 5.43: Trecho 37. Figura elaborada pela autora.

algoritmo, e aí definir quais são as métricas de sucesso. Ah, qual que é o esforço da recrutadora, horas de recrutadora por contratação de sucesso, quantas horas destinadas ali para isso, como que fazia no método atual, como que faz no novo, e roda isso por uns meses e avalia o resultado e coloca algumas métricas de sucesso. Por exemplo, essas pessoas foram razoavelmente bem avaliadas no tempo de experiência, os 90 dias. Aí tem que discutir um pouco com o pessoal de Rh, o que seria algumas self-metrics.” (Luciana)



Figura 5.44: Trecho 38. Figura elaborada pela autora.

“Porque existe aquela coisa de desempenho, de modelo é uma coisa, né? Precisão, qual a acurácia... e outra é o que é uma métrica para o negócio, que eu acho que às vezes isso é um pouco deixado de lado e aí, por isso alguns modelos entram no limbo, mas pro negócio o que é importante é que as pessoas vejam que as pessoas, se inscrevam as interações? Então isso definiria com a equipe de negócio produto.” (Renata)



Figura 5.45: Trecho 39. Figura elaborada pela autora.

“Eu imagino que a métrica de sucesso seria o algoritmo ser capaz de trazer um conjunto de pessoas adequado. Assim, uma quantidade de pessoas adequadas pra esse meu processo, (...) com uma métrica que as diferencie de maneira interessante, pra poder conduzir as partes não tão automatizadas do processo, né? Então, por exemplo, se o algoritmo me trouxer dois mil candidatos pra uma vaga, eu não vou considerar ele adequado. Vai ser muito difícil eu conduzir

entrevista com tanta pessoa. Mas, se trazer dois fulanos aí, também não é muito bom, né? Acho que essa seria uma métrica de sucesso de conjunto de pessoas.” (Denise)



Figura 5.46: Trecho 40. Figura elaborada pela autora.

– Impacto Real e Medição

Os trechos de Alice e Ana abordam a importância de medir o impacto real das implementações de ML no mundo real:

“Eu acho que essa vai ser a etapa, digamos assim, crucial, porque a gente vai conseguir calibrar exatamente para o que a gente procura, não apenas na parte automatizada, mas no que realmente a empresa está precisando ali na hora da contratação. Então, por exemplo, se a gente está fazendo essa campanha de divulgação de vaga, uma métrica importante vai ser, com certeza, quantas pessoas que se candidataram estão preenchendo os requisitos que a gente tinha definido antes, independente de ter sido pelo Machine Learning ou pela forma mais manual, digamos assim, dos processos. Então, no final das contas, tudo isso entraria ali nas métricas para avaliar o sucesso dessa campanha.” (Alice)



Figura 5.47: Trecho 41. Figura elaborada pela autora.

“E aí a pessoa vai lá e se cadastra, então talvez no caso, comparar quantas pessoas foram impactadas *versus* quantas pessoas clicaram, né? Então tipo pegar essa métrica pra quantas pessoas chegou o anúncio e elas só olharam e passaram e quantas de fato clicaram e quantas de fato se cadastraram na vaga na contratação.” (Ana)



Figura 5.48: Trecho 42. Figura elaborada pela autora.

Mariana traz um exemplo de balanceamento de vieses de gênero diante a reprodução por meio dos dados da realidade de iniquidade de gênero no meio profissional da tecnologia:

“Quando a gente treina o modelo, a gente pode usar esse tipo de informação e balancear meu dado. Suponhamos que na base que eu uso pra treinar, tem oitenta por cento de homens e vinte por cento de mulheres. Mas eu não quero chegar nesse resultado pra selecionar das pessoas. Eu não quero que cheguem oitenta por cento homens e vinte por cento mulheres. Na hora que eu for entrevistar eu quero que ele balanceie isso. Eu posso dizer olha só eu quero que a coluna que é variável gênero, fique balanceada. Então, ele vai dar um jeito, né? De conseguir balancear e, na hora de me retornar cinquenta, cinquenta. Sabe, mesmo que no aprendizado dele tinha oitenta vinte? Esse tipo de coisa é legal pensando em incorporar a diversidade, né, no retorno do algoritmo.” (Mariana)



Figura 5.49: Trecho 43. Figura elaborada pela autora.

5.7.0.5

Feedbacks e Iterações

Todos os trechos recortados ressaltam o papel da comunicação contínua e da interação entre os membros da equipe, com o cliente e com os usuários finais. A necessidade de feedback contínuo é enfatizada, tanto para ajustar o algoritmo durante o desenvolvimento quanto para avaliar sua eficácia após a implementação. O feedback dos usuários e do cliente é destacado para identificar falhas e oportunidades de melhoria.

“Eu errei, errei todos que são aderentes, não tô trazendo valor pro usuário, então acurácia você vê quanto que eu aceitei, seja positivo ou negativo, se for uma participação binária em relação ao todo. Então, se eu acertei o que não é aderente, ela vai contabilizar como acerto pra mim, porque eu tô acertando aquilo que não é válido, só que eu tô errando todos que são válidos. Então, a acurácia por si só, ela não é boa por causa disso.” (Pedro)

integração e implementação

desafios técnicos

Figura 5.50: Trecho 44. Figura elaborada pela autora.

“Ele entregou e acha que aquele modelo é ótimo, faz uma troca entre os modeladores, a gente chama de peer-review. Primeiro a gente troca, aí a gente faz uma análise assim, tem uma documentação, todas as variáveis, e o que a pessoa tá fazendo (exemplo): Ah, eu vou colocar uma individual pra cada um pontuando cada nível, aí eu vou colocar uma curva de grau dois, é uma curva de grau três. Ah, eu vou colocar uma curva e uma parte de fator, então assim eh essas decisões você tem que explicar nesse documento. A gente tem um documento inicial, a gente explica nesse documento, coloca o gráfico lá do que tá acontecendo, o Porquê você tomou essa decisão e aí vem o peer-review em cima desse documento que você fez na hora de montar o modelo, ele vai oh, não concordo com isso, eu tentei a curva de grau dois e ficou melhor ajustado ou perdeu pra mim assim ou aumenta o grau, porque a gente entende que essa variável faça parte...” (Caio)

habilidades complementares

relações dinâmicas e interrelacionais

outputs e pontos de contato comunicados claramente

busca de estímulos

comunicação do direcionamento da produção de insumos

Figura 5.51: Trecho 45. Figura elaborada pela autora.

“Ah, eu acho que tem que ser constante com o cliente e também entre o time, né, com as pessoas eu imagino, que eu vá tá desenvolvendo junto, né? Com outras pessoas, e inclusive de pessoas que foram selecionadas, né?” (Leandra)

habilidades complementares

Figura 5.52: Trecho 46. Figura elaborada pela autora.

“Ninguém pode tomar decisão unilateral, porque a gente tem um problema de viés grande, quando a gente só consegue enxergar um lado. Então, talvez a outra pessoa que tá do teu lado tá enxergando um outro, uma outra questão, um outro problema, uma outra perspectiva, então, pra mim é um default, tem que ser feito isso. A equipe tem que ser conversada. Como que a gente vai fazer isso? Que você notou e, principalmente, pra variar, o que o cliente diz? Porque como você capta a mensagem é diferente de cada um. (...) isso ajuda você ter uma visão mais ampla do que foi dito, do que foi pedido.” (Renata)



Figura 5.53: Trecho 47. Figura elaborada pela autora.

“Eu acho que o feedback precisa ser contínuo, contínuo no sentido de, ou toda semana ou de quinze em quinze dias, as pessoas externas estarem incluídas pra poder primeiro entender o que tá acontecendo e segundo pra poder palpitar, porque às vezes as pessoas que estão fora não conseguem entender o que tá acontecendo? Ou tá enviesado já, já e as pessoas de dentro já cansadas do que tão entendendo ali.” (Caio)

vieses de gênero, raça e agravantes de limitações

habilidades complementares

Figura 5.54: Trecho 48. Figura elaborada pela autora.

“Uma forma de avaliar se tá ficando bom é, ao longo do tempo, ver quanto tempo a pessoa tá na empresa. Se ela continuou, né? E seria

bom saber algumas características que talvez estejam de acordo, né? Então, a empresa contratante esperava que ela conhecesse algumas informações. Ela poderia dar um feedback pra gente, verificando ou não, talvez se essas pessoas estão correspondendo às características que eles queriam. Talvez até o funcionário, né? Ou a pessoa que foi contratada. Seria legal saber se atendeu a expectativa dele. Talvez se ele também procura uma empresa a empresa também tem que achar ele dependendo como que foi indo ao longo do tempo daria para até fazer uma paralela.” (Regina)

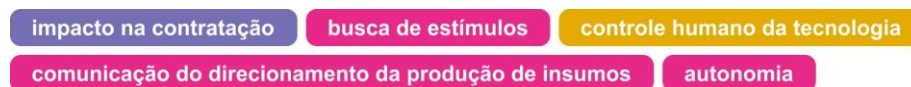


Figura 5.55: Trecho 49. Figura elaborada pela autora.

“O único jeito que você tem de perceber feedback desse tipo é através de um sistema desses de interação com o cliente. Porque quando você tá ali no teu, no teu sistema de desenvolvimento, um treinamento profissional, teve que fazer um trabalho prévio, onde você teve que fazer toda a anotação de separação, entre classe e tudo mais. Então, ali você sabe o que tá certo ou que tá errado, e você vai treinar o teu modelo baseado naquilo ali, tá? E, depois que você sai desse mundo, esse mundo tem uma variabilidade, sei lá restrita, porque você tá ali com uma população de dado um pouco restrita. Logicamente, quando você sai pro mundo, que aí a quantidade de dado é muito maior e a variabilidade dele é bem maior, também é que você começa a ter os resultados reais. E mas você já não tem aquela classificação, aquela separação, aquela anotação que você tinha ali no teu ambiente controlado, onde você fez o treinamento. Então, você precisa de alguém que te fala: é aquele caso tá falhando. Ah, esse caso tá acertando, entende? Então esse sistema de feedback é essencial, né?” (Bruno)

“A solução de feedback e interações eu acho que isso aí seriam com testes, digamos, assim como se fossem testes manuais mesmo, digamos, com pessoas, com possíveis clientes. Então, vamos dizer que eu tô fazendo uma validação. Fiz uma POC (Proof of concept), né? Eh, aí eu pego a minha POC e coloco pra vinte usuários

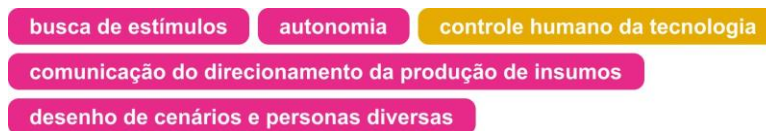


Figura 5.56: Trecho 50. Figura elaborada pela autora.

diferentes com aquele dados eles vão lá, fazem as inscrições, entram na minha rede, por exemplo, aí eu crio as vagas e eu mando pra ele essas vagas. Então a minha validação seria feita com uma quantidade de usuários, 20? Eu acho muito pouco. Talvez 50 usuários, 100 usuários tentaria mudar o máximo que eu pudesse, porque cada vez mais dados inputados (inseridos) melhor. Eu faria dessa forma, pegaria uma quantidade de pessoas pra poder utilizar o algoritmo mesmo, pra validar, porque o teste unitário fica meio um pouco viciado em alguns aspectos.” (Davi)

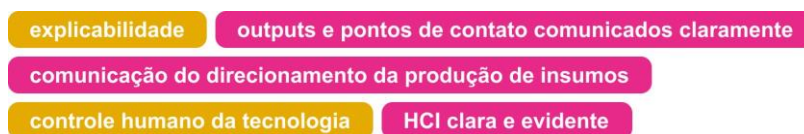


Figura 5.57: Trecho 51. Figura elaborada pela autora.

“libera (...) um beta e deixa testando. Então o pessoal vai fuçando, coloca as coisas mais mirabolantes lá, vê o que retorna, coloca o simples, vê o que retorna e aí, de acordo com o feedback, a gente vai corrigindo, né? O que tiver que isso é corrigido?” (Ana)



Figura 5.58: Trecho 52. Figura elaborada pela autora.

“Depois da primeira versão e depois de discutir com pessoas que sejam do negócio também, por exemplo, a pessoa que quer esse profissional que vai trabalhar junto com esse profissional, então chega e: Oh! O perfil que eu encontrei foi esse. Você acha que faz sentido entrevistar esse perfil? Dar uma diversificada em relação à idade em relação a gênero? Porque não necessariamente, como eu falei, pessoas com mais idade tem mais experiência do que pessoas com menos idade.” (Mariana)

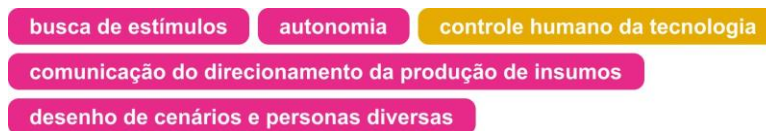


Figura 5.59: Trecho 53. Figura elaborada pela autora.

“Será que eu tô trazendo só pessoas que estão com um salário muito acima do mercado? Será que eu tô trazendo pessoas só de uma mesma empresa? (...)Ou será que eu tô trazendo com isso só um perfil, tô excluindo alguma questão de diversidade ali no meio, que teria que olhar isso? Mas, muitas vezes a gente não consegue implementar um teste AB. O mundo ideal é fazer sempre um teste AB, caso-controle em tudo, mas não necessariamente também eu vou conseguir fazer isso, porque para ser estatisticamente significativa, eu teria que contratar muita gente, ter um número de vagas grandes, então talvez eu não conseguisse fazer isso nesse mundo ideal, né? E aí algumas vezes o que a gente faz(...)já fala que é pseudo-teste, que a gente não tá fazendo mais AB (teste), faz como que funcionou antes e depois, faz o antes e depois. Só que depois eu não tô mais isolando, eu tenho a influência da variável tempo, pode ser simplesmente que chegou mais gente boa. Aí não tem nada a ver com o meu algoritmo, se eu separo aleatoriamente, tá dividido, eu consigo tirar o efeito dessa variável, dessa co-variável, mas eu não consigo tirar, mas às vezes é o que a gente tem que fazer?” (Luciana)



Figura 5.60: Trecho 54. Figura elaborada pela autora.

5.7.1

Efetividade de resultados e Impacto na contratação

Nesse momento da entrevista, os participantes foram estimulados a refletir sobre a efetividade dos resultados na aplicação de algoritmos de

inteligência artificial para a atividade de recrutamento. A efetividade, nesse caso, se refere a capacidade de produzir um efeito real sobre o processo. Dessa maneira, os desenvolvedores se colocaram a pensar conclusivamente se, de fato, adotar certas automatizações nesse processo trazem resultados reais.

Luciana fala sobre a avaliação da natureza daquilo que pode ser automatizado, ou seja, para aquelas ações que demandam decisões mais mecanizadas, a automatização pode ser efetiva.

“Então, em vez de mandar 20, manda 100. E aí tem uma visão humana em cima desses 100. Será que isso não seria um caso... Eu tenho um pouco mais de esforço, mas eu tenho uma redução de horas bem considerável entre um e outro, porque, afinal, eu não estou lá mais lendo 300 currículos, eu estou lendo 100. E talvez o resultado final, na soma, seja bom. Às vezes a gente olha muito isso, sabe? O melhor cenário, a maior automação de todas. Porque você pode fazer uma automação intermediária com inteligência artificial só para fazer uma limpeza básica, como, por exemplo, eu falei lá, uma dos requisitos era cinco anos de experiência. Será que precisa alguém ler para saber que tem cinco anos de experiência?”
(Luciana)



Figura 5.61: Trecho 55. Figura elaborada pela autora.

Ainda no contexto da natureza da avaliação, Denise fala sobre ser positiva a possibilidade de tornar o processo de seleção menos pessoal, garantindo imparcialidade através da automatização de algumas decisões.

“Assim, uma coisa que eu acho muito positiva (em relação a adotar) algoritmo, é a questão da anonimidade dele. A gente tem o viés de decisão quando você olha pra pessoas, né? Isso eu vejo muito na minha empresa pra processo seletivo interno. Por exemplo, se é uma ou outra pessoa que tá candidato e que tem o perfil parecido, ele é sujeito ao viés do contratador em relação a fatores de idade, gênero, enfim, não necessariamente a capacitação técnica é o mais decisivo. Então, se o algoritmo tiver se atendo à capacitação daquela pessoa e a características que independem desses fatores é

um impacto positivo, você pode trazer uma diversidade maior. Você pode acabar se deparando com candidatos que você não imaginaria que selecionaria.” (Denise)



Figura 5.62: Trecho 56. Figura elaborada pela autora.

Em relação ao impacto na contratação, os desenvolvedores foram estimulados a pensar sobre a possibilidade de oportunidades vedadas ou geradas para determinadas pessoas alcançadas pelo filtro do algoritmo em questão. Algumas pessoas falaram sobre vieses sociais, mesmo que ainda não tivessem trazido essa questão em momentos anteriores.

“...a gente tem que ter uma visão social do que a gente está propondo, né? Porque, realmente, estamos lidando com pessoas, com vidas, com sonhos, né? Então, eu não acho que dê para que o algoritmo faça o papel e substitua um RH com um olhar mais humano. Porque, por exemplo, eu fico pensando, temos uma pessoa que (...) estudou a vida inteira e aí na construção do currículo, talvez ela tenha perdido uma palavrinha-chave que ali tenha feito diferença na hora do algoritmo apontar essa pessoa aqui. Não necessariamente quer dizer que essa pessoa não está apta para o trabalho, né? Então, eu acho que a gente vai ter que ter um cuidado para que não tenha só esse viés mais direto. (...)Tendo esse cuidado, o impacto que isso pode gerar no processo de contratação da empresa, a princípio que eu vejo como algo positivo é a agilidade no processo que seria algo mais forte. Mas os recursos humanos, eles estão ali para que a gente lide com as pessoas. A tecnologia e os algoritmos deveriam agir apenas como uma forma de facilitador, né?” (Alice)

“Pode haver impacto e me vem na cabeça agora justamente a questão de ter que ter cuidado com os vieses, né? Na hora de fazer o algoritmo, por exemplo, (para) ele não ir pra um lado que: Ah,



Figura 5.63: Trecho 57. Figura elaborada pela autora.

são essas skills aqui, tem mais homens que realizam! Então vou selecionar mais homens. Porque os homens que são qualificados, né? Teria que ser algo pra tá atenta, né? (...) Que, enfim, não é porque tem muitos homens que se inscreveram e que colocaram ali que eles têm essa que eles sejam melhores, né?” (Leandra)



Figura 5.64: Trecho 58. Figura elaborada pela autora.

Davi e Alberto falaram sobre a realidade dos gerados pela aplicação. Davi detalha como um viés de gênero, social ou racial podem ser criados a partir da modelagem de requisitos especificamente. Alberto traz um desfecho com uma opinião sobre o papel do desenvolvedor diante as possibilidades de mitigar os efeitos gerados pelos vieses reproduzidos.

“Então, você tem que de alguma forma, ter uma quantidade de dados e conseguir inputar (inserir) os dados de uma maneira muito transparente (...)se eu pego o dado só de uma determinada região, vai dar problema porque você vai ficar enviesado. Ele fica enviesado, não tem essa coisa que fala... Ah, não existe preconceito, racismo no algoritmo! Existe sim, você cria um viés, você cria uma direção. Então, qual é a melhor maneira? Você tem que tentar aumentar, pulverizar a quantidade de dados e escalas, digamos assim vou colocar nível Brasil, ao nível nacional. Então, você tem que pegar desde o cara que estudou lá na Universidade do Amazonas, desde a menina que estudou lá no Paraná e juntar todo mundo, se esse cara se formou em Engenharia Mecatrônica, essa menina também tem uns trechos parecidos. Qual o problema com a diferença? Conseguir

ver como eles podem inputar isso na plataforma que não gere tanta diferença porque gera, tipo desigualdade, gera. Eu vou dar um exemplo assim, pega o currículo de um cara que veio lá da periferia e o currículo do cara que veio aqui da Faria Lima. O currículo do cara da Faria Lima é uma página só, nossa, talentoso e tal, o cara sabe escrever, tem boa formação, na hora de inputar o dado dele vai conseguir de uma maneira muito mais rica e se você só usar dados daquelas pessoas ali pro seu algoritmo, quando você pega o currículo do cara que vem da periferia, a similaridade não vai bater tanto e você vai começar a gerar exclusão desses candidatos. Então a melhor forma de você fazer (é escolher uma) gama de dados a nível nacional. É um trabalho mais difícil, aí tu vai ter que conseguir entender como tu pode cruzar tudo aquilo, com aquela parte de você definir como é que pega os dados, etc. E como é que eu vou cruzar esses dados que são de outras regiões, de outros lugares do Brasil como as pessoas escrevem, como as pessoas tratam isso? Aí você vai colocar tudo isso nesse banco de dados de uma maneira efetiva igual pra todo mundo. E a partir daí, que você consegue fazer um trabalho mais efetivo e mais justo pros candidatos, digamos assim, você consegue gerar mais diversidade.(...) Vai ter que fazer um fine-tuning (ajuste fino no algoritmo) porque você pode gerar um impacto em algum momento, até negativo. Como assim? Se você não fizer algo que o torne mais imparcial, pegando realmente de fato aquelas skills, os conhecimentos, você pode gerar um caminho do tipo: Ah, só vou escolher candidatos formados em universidades de São Paulo e você perde talentos de fora. É um trabalho que é interessante (...) tornar mais até popular, mais acessível, pra qualquer tipo de candidato e de igual pra igual. Só em relação à habilidade, porque tem muita gente que não importa de onde é, não importa o que faz, que é muito habilidoso, às vezes não se expressa bem, ou então não tem um bom currículo, não tem nada disso. Eu acho que a diferença vai ser como a plataforma pega os dados de uma maneira igual pra todos os candidatos, não tem sexo, não tem diferença social, o dado vai chegar igual pra todo mundo! E dali você consegue fazer um fine-tuning pra que a vaga pra aquela posição seja distribuída de fato pras pessoas que contemplam aquelas habilidades.” (Davi)

“Ah, eu acho que ativamente a gente tem que ter essa percepção e fazer sempre uma análise que considere a sub representação das



Figura 5.65: Trecho 59. Figura elaborada pela autora.

peçoas em diferentes espaços da sociedade ou esse recorte, então sei lá, pensando nesse cenário que a gente trouxe aqui. Eu acho que o imparcial não necessariamente é o melhor caminho, né? Eu acho que isso é um ponto também, né? É importante a gente ser crítico e não só imparcial. Pensando num cenário técnico completamente técnico (...) acho que os algoritmos de alguma forma têm que ser imparciais, mas a nossa leitura e o nosso posicionamento enquanto pessoa não, né? E aí, a partir desse lugar técnico que eu exerço muitas das vezes, é preciso que o meu entendimento de mundo faça com que eu crie algoritmos imparciais.” (Alberto)



Figura 5.66: Trecho 60. Figura elaborada pela autora.

5.7.2

Guia de Diretrizes para Algoritmos Sensíveis

Ao final das entrevistas, os desenvolvedores foram questionados se achariam interessante e proveitoso ter um Guia de Diretrizes para Algoritmos Sensíveis com direcionadores de ações mais inclusivas para incluir na rotina de desenvolvimento de algoritmos. Seguem três trechos em que os entrevistados mais desenvolveram sobre o questionamento.

“Eu já vi exemplos de mercado disso funcionar bem, já vi exemplos disso não funcionar bem. Tem um case que há muitos anos eu uso dando aula, que é um case da Amazon, que a Amazon fez um algoritmo de contratação, quando eu falo sobre a importância

do que entra de variável, que às vezes a gente confia muito no output (saída) do modelo e esquece que tem um input ali dentro e do histórico da base de pessoas, tem um case da Amazon sei lá, 10 anos atrás, que eu uso em aula muitas vezes, que a Amazon fez um algoritmo de contratação e eu acho que eles usavam até como variável resposta pessoas que performavam bem na empresa e tal. E no fim das contas, nenhuma mulher era selecionada nesse algoritmo. E aí, é o algoritmo que é machista? Não é o algoritmo que é machista. A história da empresa é machista e tem muito mais homem ali dentro, então o algoritmo não vai entender que quem está performando bem é homem. É isso, o algoritmo está olhando, está performando bem. Só que, no fim das contas, ele pegou a variável homem e colocou ali. Então, a gente tem que ter essa crítica. Esse caso é bem antigo e bem clássico.” (Luciana)



Figura 5.67: Trecho 61. Figura elaborada pela autora.

“Eu acho que sim, principalmente da modelagem de dados, e isso é o principal. Na minha opinião, eu vejo que quem tem que trazer esse guia numa empresa, seria gerente de produto, o PM (Product Manager) normalmente eh o quê? O engenheiro não vai desenvolver sem isso, não tem como. Ele, vai pensar em dados, como é que eu vou pegar esses dados todos e colocar no banco de dados vetorial, um exemplo. Assim, isso que ele vai pensar, a questão de como ele vai colocar, ele vai pegar tudo e colocar. Por isso que eu acho que nesse caso tem que ter um guia, não só pro engenheiro, mas principalmente pro cientista de dados. Mas isso voltado assim, cara é um projeto social, é um projeto que quer cobrir uma boa parcela da população brasileira, considerar pessoas marginalizadas e blá blá, tem que sim, tem que ter uma pessoa especializada. Porque senão, na verdade, que você vai entregar o produto que normalmente a empresa visa lucro, então você vai ter que entregar e dane-se, funcionou, se tá vendendo é bom ou se não, pula pra

outra, funciona assim. Mas se você quer garantir a assertividade, uma questão de maior pulverização, mais pessoas, etc. Tipo, eu quero conseguir que meu algoritmo alcance regiões mais pro interior do país. Esse checklist tem que ser feito na hora da modelagem dos dados, porque senão você não consegue (...) a forma que escreve é um trabalho de pesquisa, não é um trabalho assim que o cara vai definir checklist, o framework de uma vez. Ele vai ter que pesquisar o que ele quer conseguir inferir com aqueles dados de fato, como ele quer ter aquela solução, é um trabalho com design mesmo, é produto, de ver como aqueles dados podem ser imputados e fazer testes.” (Davi)



Figura 5.68: Trecho 62. Figura elaborada pela autora.

“Olha, a resposta é depende! Depende bastante, porque se os critérios forem muito bem definidos, eu acho que OK, você pode por um checklist. Eu acho importante ter um checklist genérico de coisas, por exemplo, raça, cor, essas coisas eu acho que realmente tem que ter pra ter uma equidade no processo, né? Então, pra que o algoritmo não não envie apenas a publicação só pra pessoas brancas ou só pra pessoas indígenas, ou só pra pessoas negras, e assim por diante. Então, você tem que ter um checklist de quantidade de entradas, né? Porque tem que ser algo igualitário ali e de gênero também, pra que não gere nenhum tipo de desbalanceamento. Então, os dados têm que estar balanceados para diversos grupos da sociedade. Agora, se é algo muito específico, por exemplo, é que nesse caso em específico que você me passou eu acho que o checklist genérico já solucionaria. É que eu tava falando em casos específicos que às vezes tem vieses, que são naturais. Ah, teve uma pesquisa em que constatou isso e é isso. Tem que aparecer esse viés porque é algo natural que existiu. Agora se é algo que você quer combater ou é algo que você quer, que a empresa, no caso mude o perfil e tudo mais, aí sim, pra deixar tudo igualitário. Eu acho importante

sim ter esse checklist, mas de uma forma mais genérica, de questões mais sensíveis, né!” (Vinícius)



Figura 5.69: Trecho 62. Figura elaborada pela autora.

5.7.3

Mapa mental dos marcadores

O gráfico comparativo entre as unidades de análises registradas nos trechos de cada entrevista realizada com os 20 homens especialistas e 20 mulheres especialistas pode ser visualizado na figura 5.6. Não houve peso significativo na diferença entre temas classificados. Com intuito de gerar uma



Figura 5.70: Gráfico comparativo entre unidades de análise registrada em todas as entrevistas do grupo de homens vs grupo de mulheres. Figura elaborada pela autora.



Figura 5.71: Mapa mental - Análise de resultados

síntese sobre a análise geral dos resultados extraídos das entrevistas, foi gerado um Mapa Mental com o agrupamento das unidades de análises ou marcadores na sequência em que foram sendo abordados nas narrativas de todas as entrevistas. Na figura 5.71 está representado o Mapa Mental com a síntese das categorias abordadas.

Na fase de definição de requisitos, os participantes destacaram a importância da definição das regras de negócio e do aprofundamento com os clientes, aspectos que foram associados à usabilidade e à implementação de uma pesquisa mais abrangente sobre o objetivo da aplicação. Esses elementos foram considerados fundamentais para diversificar e contextualizar melhor os diferentes cenários de uso. Alguns participantes revisitaram experiências pessoais durante essa etapa, ressaltando como vivências individuais influenciam nas decisões técnicas. Por exemplo, um participante mencionou que, durante a faculdade, não possuía um computador, o que impacta, até hoje, suas preferências na definição de requisitos, optando por outros meios em determinadas situações. Tais relatos evidenciam como a experiência de vida pode moldar as decisões técnicas atuais.

Ao abordar a utilização de dados, emergiram discussões relevantes sobre o equilíbrio entre técnicas de aprendizado de máquina e decisões humanas. Muitos participantes destacaram que, no contexto de algoritmos de recomendação, o aprendizado por similaridade pode gerar preocupações, uma vez que o algoritmo tende a reproduzir padrões presentes nos dados. Este ponto foi

identificado como um momento sensível no processo de desenvolvimento, especialmente quando se trata da seleção de critérios, onde questões de gênero e diversidade foram levantadas. Alguns especialistas enfatizaram a predominância de homens em certos conjuntos de dados e discutiram maneiras de aumentar a visibilidade e tornar os anúncios mais inclusivos.

No que diz respeito ao desenvolvimento, muitos participantes enfrentaram desafios relacionados à obtenção de dados, destacando a diversidade metodológica e as limitações técnicas. Embora diversos métodos tenham sido considerados, a viabilidade de processamento e a capacidade de extrair dados de forma eficaz foram frequentemente citadas como barreiras. Além disso, as questões de customização e precisão foram discutidas no contexto de como expandir a abrangência da aplicação, sempre com foco no impacto real da utilização.

Por fim, durante a fase de validação de métricas, as discussões foram centradas em três categorias principais, com ênfase no desenvolvimento de métricas de sucesso baseadas no conhecimento do negócio. A multidisciplinaridade foi apontada como essencial, com a participação de clientes e profissionais de diferentes áreas, sendo considerada muito relevante para a definição e avaliação das soluções propostas. Questões de aplicabilidade e transparência também foram amplamente discutidas, com preocupações sobre como garantir a visibilidade de gatilhos e a comparabilidade dos resultados ao longo do tempo. Muitos participantes relataram experiências pessoais relacionadas ao impacto social direto dos algoritmos, trazendo reflexões sobre gênero, diversidade e o equilíbrio entre decisões automatizadas e humanas.

6

Discussão do Guia de Diretrizes para Algoritmos Sensíveis.

6.1

Discussão por indicadores

No decorrer da realização das entrevistas, muito foi falado pelos desenvolvedores sobre a necessidade e importância em trocar impressões e debater questões técnicas e não técnicas com seus pares e sobretudo com pessoas especialistas em temas mais diversos. Nesse contexto, observa-se que essa necessidade declarada também carrega a aspiração por dividir responsabilidades, ou seja, ter com quem dividir o peso de algumas decisões que podem gerar maiores consequências, tanto para as aplicações desenvolvidas quanto para seus usuários. Desse modo, entendendo que no recorte específico dessa pesquisa, a relação de alcance e consequências sociais como efeito da interação de pessoas (usuários) com sistemas (algoritmos) se revela, chamamos então, os algoritmos em questão de **algoritmos sensíveis**. Nesse contexto, houve a necessidade de discussões mais aprofundadas sobre esses momentos importantes na tomada de decisão, e a ideia da aplicação de um Guia de Diretrizes para Algoritmos Sensíveis foi bem recebida. A partir da visão da tecnologia como produto, envolvemos os usuários e desenvolvedores como a parte humana desse eixo, reforçando a ideia do **algoritmo centrado no humano**. Nesse contexto, o Guia com Diretrizes para Algoritmos Inclusivo, traz as diretrizes como orientações e princípios que especialistas em IA devem seguir para criar algoritmos que respeitem e promovam a inclusão e a diversidade.

Portanto, com base na interpretação do checklist de Kat Holmes (Tessier, 2021) foi apresentado os direcionadores do **Guia com Diretrizes para Algoritmos Sensíveis** para serem discutidos um a um no grupo de foco. Esta foi a estratégia para provocar uma reflexão mais profunda sobre as práticas atuais de desenvolvimento de produtos e transpor para a realidade dos desenvolvedores, as necessidades envolvidas no processo de modelagem dos algoritmos. Esse checklist é reconhecido por seu enfoque na inclusão, ao sugerir uma série de diretrizes que promovem o desenvolvimento de produtos que contemplam a diversidade dos usuários. A pesquisadora propõe que este conjunto de orientações seja usado como uma ferramenta para desafiar as

abordagens convencionais, principalmente as metodologias ágeis. Estas, por sua natureza de agilidade e foco na entrega rápida, muitas vezes, simplificam ou ignoram variações e necessidades específicas cruciais para a inclusão efetiva de todos os usuários. O uso do checklist de Holmes busca expandir o entendimento e a empatia dos desenvolvedores em relação às várias realidades enfrentadas pelos usuários, incentivando a criação de soluções tecnológicas mais acessíveis e equitativas. A discussão desses direcionadores tem como norte, reforçar a responsabilidade social dos desenvolvedores na construção de tecnologias inclusivas.

A adaptação do checklist para o **Guia de Diretrizes para Algoritmos Sensíveis** fica definida com os seguintes direcionadores observados na figura 6.1.

Indicadores – Design Justice

- 1 Considere a **diversidade sob a perspectiva das interações humanas** e considere a maneira como as pessoas mudam ao longo do tempo
- 2 **Identifique vieses relacionados à capacidade, gênero, raça e/ou acessos** e preveja possíveis interações incorretas que possam acontecer
- 3 **Crie uma diversidade** de caminhos para a **participação em sua experiência**
- 4 Assuma que o projeto **terá incertezas, necessidade de iterações e adaptações** e trabalhe com a utilização de habilidades complementares
- 5 Desenvolva uma **alfabetização básica em acessibilidade** e se aprofunde nos critérios de acessibilidade específicos relevantes para suas soluções
- 6 Adote uma **definição mais flexível de projeto**. Abra os processos e convide pessoas com habilidades relevantes a contribuir com seu projeto, mesmo que não sejam programadores (as)

Figura 6.1: Guia de Diretrizes para Algoritmos Sensíveis

Foram realizados dois grupos de foco com duração média de 1h30min. O primeiro grupo foi realizado com 07 desenvolvedoras mulheres e o segundo grupo foi realizado com 06 desenvolvedores homens, ambos entrevistados na primeira fase. Todos os participantes foram convidados por serem bem participativos nas entrevistas. A decisão de fazer grupos separados por gêneros, em essência, teve como intenção deixar os participantes mais confortáveis para colocarem suas particularidades, que podem derivar da experiência vivida pelo gênero. O perfil dos participantes está na tabela 6.1 e 6.2.

Tabela 6.1: Perfil dos participantes: grupo de foco com mulheres

Identificação	Formação Acadêmica	Idade	Experiência em modelagem algorítmica	Cargo
Ana	MBA em Ciência de Dados e Big Data	35-44	Experiência Significativa (Uso frequente)	Consultora Pleno em Analytics and Data
Leandra	Graduação em Sistemas Computacionais	25-34	Pouca Experiência (uso ocasional)	Consultora Pleno em Analytics and Data
Renata	Mestrado em Ciência da Computação	25-34	Experiência Significativa (Uso frequente)	Consultora Senior em Analytics and Data
Mariana	Graduação em Estatística e Mestrado em Ciência de Dados	25-34	Especialista (ampla experiência / liderança)	Consultora Senior em Analytics and Data
Alice	Graduação em geofísica e Mestrado em Dinâmica dos Oceanos e da Terra	25-34	Alguma experiência (conhecimento básico)	Consultora Pleno em Analytics and Data
Regina	Graduação em Engenharia Química	25-34	Experiência Significativa (Uso frequente)	Consultora Pleno em Analytics and Data
Rebeca	Graduação em estatística e MBA em Big Data e Business Analytics	35-44	Experiência Significativa (Uso frequente)	Consultora Pleno em Analytics and Data

Tabela 6.2: Perfil dos participantes: grupo de foco com homens

Identificação	Formação Acadêmica	Idade	Experiência em modelagem algorítmica	Cargo
Caio	Graduação em Estatística	25-34	Especialista (ampla experiência / liderança)	Consultor em Pricing & Analytics
Francisco	Graduação em Engenharia Elétrica	35-44	Experiência Significativa (uso frequente)	Machine Learning Engineer
Davi	Graduação em Ciência da Computação incompleta	25-34	Experiência Significativa (uso frequente)	Engenheiro de Software Senior
Bruno	Doutorado em Materials Science and Engineering	35-44	Especialista (ampla experiência / liderança)	Gerente em AI & Data
Renato	Doutorado em Computação	35-44	Especialista (ampla experiência / liderança)	Gerente em AI & Data
Afonso	Graduação em Engenharia	25-34	Experiência significativa (uso frequente)	Consultor Senior em Analytics and Data

Durante a dinâmica, os participantes foram direcionados para seus pai-

néis virtuais individuais localizados em um quadro da plataforma Miro. Cada

direcionador do Guia foi lido pela mediadora, e na sequência, os participantes tinham dois minutos para refletirem sobre o que entenderam e se possível, responderem as seguintes perguntas que auxiliavam a trazer para suas realidades de trabalho:

- O que você entendeu sobre esse item?
- Acha realizável no dia a dia?
- Como você aplicaria? Quais meios buscaria?
- Esse Item te fez pensar em algo que não costuma pensar durante o desenvolvimento de algoritmos?

O painel com as perguntas e os direcionadores foi disponibilizado a todos os participantes que podiam acessá-lo de forma colaborativa e síncrona, na figura 6.2 observamos como o painel virtual estava organizado.

Item	Descrição	O que você entendeu sobre esse item?	Acha realizável no dia a dia?	Como você aplicaria? Quais meios buscaria?	Esse item te fez pensar em algo que não costuma pensar durante o desenvolvimento de algoritmos?
01	Considere a diversidade sob a perspectiva das interações humanas e considere a maneira como as pessoas mudam ao longo do tempo	Diversidade de gêneros, neurodiversidade, de corpos, de raça, de deficiências, de idade.	sim, mas é preciso engajar e pensar nas necessidades.	no caso de aplicar, buscando opções que engajem a diversidade.	me fez pensar na minha própria experiência nas interações humanas, como mulher e não muito ágil, experimentando insegurança, na dificuldade de fazer praqu岸er minha voz e opiniões.
02	Identifique vieses relacionados à capacidade, gênero, raça e/ou acessos e preveja possíveis interações incorretas que possam acontecer	Saber prever possíveis vieses que possa vir a sofrer para poder pensar para a diversidade em si.	sim, mas não como um todo, talvez não seja possível se lidar de todos os vieses.	buscando mapear possíveis interações incorretas, com testes rigorosos e feedback.	sim, ao 1.
03	Crie uma diversidade de caminhos para a participação em sua experiência	é preciso pensar nas diversidades e nos aspectos que possam surgir no meio do caminho da experiência de um usuário, seja no caso de dados ou em si mesmo, se de algo que já foi desenvolvido e se lidar com dificuldades com tecnologia, por exemplo.	sim, mas nem tudo, algo muito subjetivo também é muito complexo, mas com certeza muitos cenários são possíveis de serem implementados.	mapeando as necessidades para produtos, desenvolvendo caminhos possíveis e realizando a verificação.	Esse item te fez pensar em algo que não costuma pensar durante o desenvolvimento de algoritmos?
04	Assuma que o projeto terá incertezas, necessidade de iterações e adaptações e trabalhe com a utilização de habilidades complementares	Assumir habilidades complementares como aquelas além do que as habilidades profissionais nos permitem, por exemplo, se não sou hábil, então, saber lidar com frustrações, adaptações, ter paciência e estar aberto em soluções alternativas para problemas.	sim, é algo que mexer habilidades que não são "técnicas" nos cursos, já que muito de gênero, mulheres que não habilidades mais encontradas nos homens.	multa reflexão e paciência para poder pensar em possíveis soluções.	Esse item te fez pensar em algo que não costuma pensar durante o desenvolvimento de algoritmos?
05	Desenvolva uma alfabetização básica em acessibilidade e se aprofunde nos critérios de acessibilidade específicos relevantes para suas soluções	não sei se entendi bem, talvez mudar as posturas, entendendo que podem acontecer o usuário, pensando na meu público específico, meu público é de adultos, talvez entender quais são as necessidades que um adulto com necessidades poderia encontrar.	sim, mas requer um tempo de tempo e talvez não da empresa de ser inclusiva com o cliente.	buscando meios de melhorar a jornada dos usuários, em busca das melhorias para melhorá-los.	Esse item te fez pensar em algo que não costuma pensar durante o desenvolvimento de algoritmos?
06	Adote uma definição mais flexível de projeto. Abra os processos e convide pessoas com habilidades relevantes a contribuir com seu projeto, mesmo que não sejam programadores(as)	sobre a heterogeneidade da equipe, a importância de ter um diversidade, pois é daí que nasce a pluralidade de ideias.	sim, mas não apenas programando das habilidades e sim de uma iniciativa da empresa / grãda.	criar alternativas para cargos fixos na empresa, buscar de manter e apoiar de pessoas diversas.	Esse item te fez pensar em algo que não costuma pensar durante o desenvolvimento de algoritmos?

Figura 6.2: Painel para preenchimento dos questionamentos debatidos no grupo de foco acerca de cada direcionador do Guia.

Após o término do tempo destinado ao aprofundamento, a mediadora chamava todos ao debate do item, com um tempo reservado em média de 10 minutos para cada direcionador.

Neste capítulo, serão apresentadas as principais percepções e ideias relevantes emergentes dos debates divididos por direcionadores com foco em discutir diretrizes ou recomendações para o aprimoramento de práticas de desenvolvimento baseadas em uma compreensão mais profunda dos modelos mentais que promovam equidade e eficácia.

Também iremos comparar os modelos mentais entre grupos de desenvolvedores de diferentes gêneros (feminino e masculino), para entender variações que possam influenciar a abordagem ao desenvolvimento algorítmico.

6.1.1

1. Considere a diversidade sob a perspectiva das interações humanas e considere a maneira como as pessoas mudam ao longo do tempo

Grupo de foco com mulheres

A maioria das participantes (Renata, Leandra, Rebeca, Regina), introduziu o debate trazendo a **diversidade** do ponto de vista interno da equipe de trabalho e a relação de **tempo** se referindo as mudanças ocasionadas dentro da equipe. Na sequência, todas concluíram a necessidade de promoverem a modificação interna, na expansão de um olhar para as outras diferenças, para questões que estão na sociedade.

Com a visão do desenvolvimento do algoritmo como produto, Renata trouxe a dificuldade dos dados refletirem genuinamente os perfis e especificidades dos usuários. Renata discute sobre os limites dos algoritmos em prever perfis de usuários, destacando a complexidade de lidar com padrões de comportamento que são inerentemente fluidos e sujeitos a mudanças frequentes. A dependência de dados recentes, como em algoritmos de recomendação, pode não capturar a essência completa de comportamentos ou intenções futuras, o que é um desafio particular em aplicações como prevenção de fraudes ou personalização de serviços. Outro debate decorrente do entendimento da diversidade trazida nesse primeiro indicador, é o impacto positivo da diversidade no desenvolvimento pessoal. Segundo as participantes, estar em um ambiente diversificado não só enriquece as interações, mas também promove um crescimento pessoal mais profundo. A exposição a diversas perspectivas e experiências pode ampliar a visão de mundo dos indivíduos e sua capacidade de adaptação. Regina fala sobre a experiência de ser mulher em um ambiente de tecnologia e Ana traz questões do dia a dia que a fazem enxergar as maneiras diversas de existir em sociedade:

"Que nem o time que eu estou hoje, a maioria era homem. Começou só comigo ali de mulher. Então, você vai vendo que no começo eles não queriam te ouvir. Querendo ou não, acaba sendo machista. Tem uma

certa trava, mas vai mudando. Você vai conquistando ali seu espaço e eles acabam até baixando a guarda."(Regina)

"A gente tem que pensar nisso e ir se rodeando disso no dia a dia, né? Então, acho que quando a gente só convive com os nossos iguais. A gente meio que para de pensar. Que nem o exemplo do poste no meio da rampa de acesso. Então, assim. Eu tenho um filho. Então, eu sei o que é. Quais são as dificuldades ali de uma criança. Porque eu vejo pelos olhos dele muitas coisas no mundo. Meus pais são idosos. Então, eu vejo pelos olhos deles muitas coisas no mundo. E eu acho que é dessa forma que a gente vai vendo. Eu tenho um cunhado que ele tem deficiência. Então, eu vejo muita coisa pelos olhos dele. E acho que isso vai contribuindo também, né?"(Ana)

A partir dessas reflexões sobre a diversidade no dia a dia, Ana abordou a importância de considerar uma ampla variedade de diversidades ao desenvolver produtos, como gênero, neuro diversidade, tipos de corpos, raça, idade e deficiências. Ela exemplificou isso ao falar sobre a necessidade de adaptar a usabilidade de aplicativos para atender a públicos específicos, como idosos, que podem enfrentar mais dificuldades com tecnologia, ou crianças, que utilizam aplicativos educativos. Ana enfatizou a importância de compreender e interpretar as necessidades do público alvo, incluindo pessoas neuro diversas e daltônicas, para garantir que os produtos sejam acessíveis e úteis para todos. Até esse momento, e mesmo durante as entrevistas, os participantes ainda não haviam levantado questões sobre acessibilidade.

Na conclusão desse primeiro direcionador, Alice conclui aprofundando a reflexão sobre mais do que incluir diversidade em seu dia a dia, é necessário atuar ativamente para mudar e ampliar a consciência para ter mais confiança em agir ao encontro da inclusão da diversidade:

"Então, eu acho que é possível a gente lidar com potencialidades, com essas mudanças, desde que a gente reflita sobre onde a gente quer inserir isso. Se é só num algoritmo ou se é na nossa vida, mas eu acho que tudo isso passa por um caminho de reflexão, de percepção como indivíduos. Não acho que é algo que seja construído só por osmose. Acredito que a gente tem que ter uma reflexão mais profunda e realmente refletir onde a gente quer chegar. Então, se a gente talvez está querendo construir um algoritmo, a gente tem que refletir no impacto disso e em como nós queremos que ele seja. Então, tem que pegar o barco, o remo e nós mesmos fazermos esse processo. Que nem sempre é algo simples."(Alice)

Grupo de foco com homens

De modo geral, os participantes trouxeram exemplos práticos de como devem considerar sobretudo a passagem de tempo para suas aplicações e produtos. A análise das contribuições destaca a necessidade de considerar a diversidade e a evolução temporal das informações pessoais, especialmente aqueles baseados em dados.

Davi traz uma perspectiva prática sobre a validação de identidade, ressaltando que documentos antigos podem mostrar uma aparência muito diferente da atual devido às mudanças ao longo do tempo. Esta observação sublinha a importância de considerar essas mudanças para garantir a precisão na identificação de indivíduos.

Renato, por sua vez, discute a variabilidade dos dados pessoais, como nome e gênero, que podem mudar ao longo do tempo. Ele destaca a importância de adaptar algoritmos para considerar essas mudanças, garantindo que os dados reflitam a realidade atual do usuário e as recomendações sejam relevantes e precisas.

Caio foca na indústria de seguros, onde a precificação é diretamente afetada por mudanças como a idade do condutor ou do veículo, além de eventos globais como a pandemia de COVID-19. Ele aponta para a necessidade de controles no desenvolvimento de modelos para minimizar a influência de variáveis temporais e pessoais, assegurando que o risco seja calculado corretamente.

Afonso, trabalhando em um contexto de manutenção e operações, percebe a influência da demografia do usuário em soluções tecnológicas. Ele observa que um público predominantemente masculino e conservador pode não estar aberto a novas implementações, o que requer uma compreensão detalhada do público-alvo para melhor engajamento e aceitação das inovações.

Finalmente, Francisco fala sobre produtos financeiros, destacando a necessidade de explicar mudanças nos dados dos usuários, como uma mudança de gênero, às autoridades regulatórias. Ele ressalta a importância de ter sistemas capazes de acompanhar e justificar alterações nos dados para assegurar a conformidade legal e a equidade no tratamento dos clientes.

Essas contribuições derivadas do primeiro direcionador demonstram a complexidade de trabalhar com dados pessoais em ambientes dinâmicos, onde a capacidade de adaptar-se a mudanças contínuas e de entender profundamente o contexto do usuário é crucial para o desenvolvimento de soluções eficazes e justas.

6.1.2

2. Identifique vieses relacionados à capacidade, gênero, raça e/ou acessos e preveja possíveis interações incorretas que possam acontecer

Grupo de foco com mulheres

Renata destaca a complexidade e a inevitabilidade dos vieses nos dados e nos modelos de inteligência artificial. Ela aponta que o mundo real é intrinsecamente enviesado e complexo, e isso se reflete nos dados que usamos para treinar modelos. Renata chama a atenção para a responsabilidade dos desenvolvedores em questionar continuamente seus próprios preconceitos e os dos dados que utilizam, sublinhando a importância de não aceitar acriticamente os dados como um reflexo fiel da realidade.

Mariana complementa a discussão de Renata, trazendo a perspectiva de que as informações pessoais, como gênero e idade, são frequentemente cruciais para a precisão dos algoritmos, mas devem ser usadas com cautela. Ela fala sobre a necessidade de justificar claramente a inclusão de tais dados para evitar injustiças, especialmente em aplicações sensíveis como seguros de automóveis, onde dados estatísticos mostram diferenças significativas de comportamento entre diferentes grupos. Mariana analisa esse segundo direcionador em suma pela necessidade incitada de levantar questões inclusivas, a depender do produto, como destacado no seguinte trecho:

"Vamos falar, por exemplo, de pessoas com deficiência auditiva, precisam muitas vezes de legenda ou ela precisa de um intérprete ali para poder estar auxiliando em alguma determinada informação. Muitos sites já pensam nisso. O YouTube tem legenda automática, cursos online estão ali tentando incluir essas pessoas também. Mas, ao mesmo tempo, é muito comum da gente ver que até você ter uma pessoa com deficiência auditiva no teu curso, você não pensou em desenvolver uma legenda automática, por exemplo. Até você ter um cadeirante como um consumidor do seu produto final, você não pensou que era importante ter a rampa. Então, é muito complexo. Eu estou aprendendo com vocês" (Mariana)

Alice reflete sobre a importância de abordar os projetos tecnológicos com uma "ótica sensível", sugerindo que a sensibilidade não se limite apenas ao tratamento de dados, mas se estenda a todo o processo de desenvolvimento. Ela critica a tendência de considerar a diversidade apenas quando é financeiramente conveniente e defende que a inclusão e a consideração de diferentes perspectivas sejam uma prioridade desde a concepção até a finalização dos projetos.

Regina e Rebeca trazem exemplos práticos que ilustram como os vieses podem se manifestar e ser perpetuados através da tecnologia e nas práticas de

contratação. Regina discute como preconceitos de idade e estado civil podem influenciar as decisões de contratação, enquanto Rebeca enfatiza a importância de fazer pesquisas e envolver pessoas com experiências relevantes antes de desenvolver soluções tecnológicas, especialmente quando se trata de atender a necessidades de grupos específicos, como pessoas com deficiência.

Em suma, as discussões refletem uma consciência crescente sobre os desafios éticos na tecnologia e a necessidade de uma abordagem mais holística e responsável no design de sistemas inteligentes.

Grupo de foco com homens

Na discussão do segundo direcionador, os participantes trazem exemplos práticos e confirmam uma visão de que as questões de vieses de gênero, raciais e sociais gerados pelos algoritmos estão em evidência, fazendo com que necessariamente as empresas que vendem aplicações de IA, tenham que identificar e pensar em soluções efetivas para esses casos. Afonso, Caio e Flavio, convergem em torno da importância de reconhecer e abordar vieses em dados e algoritmos, bem como as percepções sociais que podem influenciar na aceitação ou não de tecnologias e serviços que apresentem vieses. Afonso relata sua experiência direta e real com a desigualdade de tratamento em situações profissionais baseadas no gênero, em que clientes se demonstram mais confiantes se tiver um homem respaldando as decisões técnicas aplicadas. Este exemplo ressalta a relevância de considerar essas dinâmicas sociais ao dirigir reuniões e projetos, especialmente quando se trata de equidade de gênero. Caio discute como preconceitos podem distorcer a interpretação de dados, especificamente em relação a estereótipos de gênero. Ele ressalta que a realidade dos dados frequentemente contradiz esses estereótipos, como demonstrado em estudos sobre a condução por homens e mulheres.

Caio enfatiza a necessidade de formular hipóteses e verificar os dados de forma objetiva para assegurar decisões justas e baseadas em evidências. Flavio aborda os riscos associados a vieses em algoritmos de reconhecimento facial, citando estudos que identificam altas taxas de erros, especialmente em desvantagem para indivíduos negros e mulheres. Ele destaca a necessidade de uma análise cuidadosa dos dados e resultados dos modelos para evitar perpetuar injustiças e garantir que as tecnologias sejam desenvolvidas e aplicadas de maneira ética e justa.

A discussão para este direcionador sugere que, independentemente do contexto específico — seja na interação direta com clientes ou no desenvolvimento de tecnologias de dados —, é crucial manter uma vigilância constante contra vieses e trabalhar proativamente para mitigar seus efeitos, assegurando que a tecnologia e as práticas de negócios promovam a equidade e a justiça.

De modo geral, foi consenso que esse direcionador se confunde com o primeiro, parecendo trazer um contexto redundante.

6.1.3

3. Crie uma diversidade de caminhos para a participação em sua experiência

Grupo de foco com mulheres

A discussão central para esse direcionador alcançou as visões sobre a participação dos usuários e de líderes no desenvolvimento, em como suas percepções podem voltar para o projeto e contribuir para a criação de experiências mais ajustadas (ou forçadas) para determinados perfis.

Regina aborda a necessidade de incluir ativamente os consumidores no processo de desenvolvimento para focar nas "dores" e necessidades reais dos usuários, sublinhando a importância de acompanhar continuamente o feedback ao longo da vida do projeto. Ela destaca que entender as dores do processo não é apenas teórico; é crucial ver que informações realmente importam na prática, onde a interação do usuário com o produto pode fornecer percepções para melhorar a qualidade do produto final.

Renata continua o tema, focando na relevância das métricas na avaliação do sucesso de um produto. Ela critica a prática comum de se concentrar unicamente em métricas como vendas ou número de usuários, argumentando que métricas de engajamento e satisfação são igualmente importantes. Renata sugere a criação de métricas mais personalizadas e diversificadas que possam realmente avaliar o impacto e a utilidade de um produto para seu público. Ela também ressalta a importância de avaliar algoritmos preditivos além de suas métricas de sucesso imediatas, questionando os resultados que eles propõem e como esses resultados afetam a sociedade.

Rebeca aponta que as diretrizes do projeto, influenciadas pela liderança, podem afetar significativamente como os objetivos são alcançados, incluindo a tendência de descartar itens prematuramente para atender a metas quantitativas. Esse exemplo ilustra como a liderança pode impactar não apenas a execução do projeto, mas também a integridade e a eficácia dos resultados.

Grupo de foco com homens

Este foi o direcionador em que os participantes disseram que mais os estimularam a pensar sob perspectivas diferentes.

Davi ressalta a importância de considerar a diversidade de participantes em experimentos ou testes de produtos, sugerindo que compreender e mitigar as barreiras enfrentadas por diferentes grupos pode contribuir para a inclusão e a relevância dos resultados dos testes. Essa abordagem é uma direção

para garantir que os produtos sejam acessíveis, independentemente de suas circunstâncias particulares.

Renato aborda a personalização de produtos de uma perspectiva prática para falar das falhas nas configurações personalizadas usando o exemplo da aracnofobia de sua esposa, em que ainda que configure as plataformas para bloquear aparecem imagens de aranhas. Ele sugere que permitir aos usuários especificar suas preferências e aversões pode melhorar a experiência do usuário, e deve-se investir nisso, embora reconheça que isso pode complicar o desenvolvimento e aumentar os custos.

Afonso discute a importância do design da interface do usuário, apontando como a acessibilidade de um produto pode ser limitada ou facilitada pelo modo como é apresentado. Ele observa que simplificar o acesso a ferramentas tecnológicas pode melhorar a experiência de uso, especialmente para usuários que podem ser menos tecnicamente proficientes ou relutantes em adotar novas tecnologias.

"Francisco destaca a importância da diversidade de experiências durante o desenvolvimento do produto. Conviver com pessoas de diferentes habilidades e necessidades pode revelar deficiências no design que, de outra forma, poderiam passar despercebidas. Ele ressalta como interações diárias com pessoas que têm diferentes capacidades podem contribuir para evitar erros e melhorar a funcionalidade dos produtos:

"tenho um amigo que tem deficiência visual e, convivendo com ele, eu percebi tanta coisa, (..) exemplo, pra mim é ótimo chegar no meu prédio e ter três elevadores, apertar um botão e chamar qualquer um dos três, mas para ele isso aí é um problemão, porque ele não sabe onde vai chegar o elevador dele. (...) realmente você não tendo contato com o diferente, a chance de você fazer um produto também inadequado mesmo para uma pessoa que tem uma questão de capacidade, né, mas estar próximo é uma ajuda, é muito importante para a gente diminuir o erro e abrir as perspectivas também."

Finalmente, Bruno reflete sobre a tensão entre atender a um nicho específico *versus* abordar uma diversidade mais ampla de necessidades. Ele sugere que, dependendo dos objetivos do projeto, pode ser benéfico "falhar rápido" e adaptar-se com base no feedback do usuário, em vez de tentar antecipar todas as possíveis necessidades antes do lançamento.

6.1.4

4.Assuma que o projeto terá incertezas, necessidade de iterações e adaptações e trabalhe com a utilização de habilidades complementares

Grupo de foco com mulheres

Renata mergulha na questão da gestão de incertezas em projetos de desenvolvimento de software. Ela critica a tendência das equipes de planejar para um "mundo ideal" onde tudo funciona perfeitamente, sem considerar os inevitáveis imprevistos que caracterizam a maioria das rotinas de projeto. Renata enfatiza a importância de preparar-se para o "pior caso" em vez de apenas esperar pelo melhor, sugerindo que uma abordagem mais pessimista ao planejamento pode, na verdade, levar a resultados mais realistas e gerenciáveis. Esta perspectiva ressalta a necessidade de um planejamento flexível que possa se adaptar a circunstâncias inesperadas, melhorando assim a gestão de expectativas do cliente e a entrega de projetos.

Ana interpretou esse direcionador de maneira a ressaltar as *soft skills* que são demandadas ao trabalhar com desenvolvimento de produtos de tecnologia, destacando a capacidade de gerenciar a paciência e os humores em uma equipe, especialmente quando se lida com incertezas. Ela discute o papel das emoções no local de trabalho e como estereótipos de gênero podem influenciar a percepção das habilidades de gestão. Ana sugere que as mulheres, frequentemente vistas como mais aptas em habilidades interpessoais, podem ser valorizadas em posições gerenciais devido a essas competências percebidas. Essa discussão levantou questões sobre como os preconceitos e as expectativas de gênero moldam as trajetórias profissionais e as dinâmicas de equipe.

Mariana já interpretou pelo lado do produto em si, afirmando que é preciso considerar e prever as possíveis adaptações. Nessa linha, ela foca na importância da integração de habilidades complementares entre as equipes e afirma que considera importante envolver pessoas 'não técnicas' para poderem ter a experiência do produto e retornar com feedbacks. Assim, ela acredita que os produtos finais sejam verdadeiramente úteis e representativos das necessidades do mercado.

Regina complementa a discussão sobre flexibilidade e abertura para novas abordagens, alertando contra o risco de se apegar rigidamente a métodos conhecidos sem considerar as necessidades específicas de diferentes projetos ou clientes. Ela fala sobre a importância de avaliar cada situação individualmente e estar aberto a alterar as estratégias conforme necessário.

Grupo de foco com homens

Bruno puxa a discussão defendendo a ideia de que "falhar o quanto

antes é a melhor coisa possível"no contexto de desenvolvimento de software. Esta é uma referência ao conceito de "falha rápida" adotado em métodos ágeis, que enfatiza a importância de identificar rapidamente deficiências para iterar e melhorar o produto. Este entendimento reconhece a natureza iterativa e adaptativa do desenvolvimento moderno de software, que busca minimizar o risco e maximizar a relevância do produto final.

Na sequência, Caio discute as complexidades nas interações entre áreas pares que não possuem o conhecimento técnico. Ele menciona o uso de "provas de conceito" como uma ferramenta para alinhar expectativas e validar requisitos já desde o começo do desenvolvimento, segundo ele mesmo em termos de estruturação de modelagem de software. Essa estratégia visa facilitar a compreensão mútua e garantir que a entrega final esteja alinhada com as necessidades do cliente. Esta parte do debate enfatiza a realidade de ambientes corporativos, onde a comunicação eficaz entre diferentes partes interessadas pode ser tão crítica quanto o próprio desenvolvimento técnico.

Por fim, Francisco traz uma nova perspectiva, sobre as oportunidades geradas através das incertezas. Quando não há algo definido, há margem de criação e se isso for feito em conjunto com o cliente, há a possibilidade de entender de fato as reais necessidades a serem atingidas. Ele destaca a utilidade de dados e experimentos para moldar o produto final em conjunto com os clientes, uma prática que ressalta a co-criação e a adaptabilidade.

6.1.5

5. Desenvolva uma alfabetização básica em acessibilidade e se aprofunde nos critérios de acessibilidade específicos relevantes para suas soluções

Grupo de foco com mulheres

Ana inicia o debate declarando a insegurança no entendimento, mas sugere estudar as possíveis adversidades que podem acometer o usuário, mirando público específico em questão. Ela afirma que a melhor maneira de entender as necessidades específicas, será buscando meios de mapear a jornada dos usuários em busca dos empecilhos para resolvê-los.

Leandra, Regina, Renata e Alice, trazem esse direcionador para um contexto mais prático, pensando na utilização de guias, manuais para escolher tamanho de fontes, cores, forma de escrever, resumindo em compreender ações para promover uma alfabetização em acessibilidade. Destacam a importância de praticar essas ações desde o início do desenvolvimento do projeto, e de maneira sistêmica, fazer a documentação para ser utilizada em outros projetos.

Mariana destaca que é importante buscar soluções para minimizar o impacto das pessoas com deficiência principalmente.

Grupo de foco com homens

Afonso observa que, em sua experiência, a inclusão de características de acessibilidade muitas vezes não é considerada. Afonso sugere que a tendência de negligenciar a acessibilidade acontece, pois não é um quesito considerado como padrão.

Por outro lado, Caio descreve uma abordagem mais proativa e inclusiva dentro de sua organização, onde há grupos dedicados a cuidar de necessidades específicas de acessibilidade, como idosos e pessoas com deficiências. Ele menciona reuniões trimestrais onde experiências são compartilhadas e soluções para problemas de acessibilidade são discutidas e implementadas.

O debate revelou dois modelos distintos de abordar a acessibilidade: um modelo mais reativo e limitado, mencionado por Afonso, e um modelo proativo e integrativo, exemplificado pela experiência de Caio. A diferença entre essas abordagens traz o questionamento do momento em que se aborda considerações de acessibilidade, se desde o início do processo de desenvolvimento e não apenas como uma resposta a requisitos específicos de um subconjunto de usuários, mas sim, como uma prática padrão que beneficia todos os usuários.

6.1.6

6. Adote uma definição mais flexível de projeto. Abra os processos e convide pessoas com habilidades relevantes a contribuir com seu projeto, mesmo que não sejam programadores(as).

Grupo de foco com mulheres

Houve o consenso entre as participantes que o direcionador impulsiona a participação multidisciplinar no projeto, sendo uma etapa durante a concepção para abrir o projeto para contribuição de outras pessoas com diferentes habilidades, inclusive as não técnicas.

Alice fala que esse direcionador a faz pensar sobre a importância de se permitir "fora da caixinha" e destaca a importância de iterar a partir de feedbacks e registrar as vantagens e dificuldades desse formato de processo mais flexível.

Regina adiciona a diversidade nas soluções, ou seja, olhar para outras maneiras de resolver um mesmo problema, buscando expandir. E também, pensa que as pessoas com habilidades diversas envolvidas no processo, precisam também representar diversidade, sendo mulheres, idosos, por exemplo.

Na mesma linha, Ana destaca a heterogeneidade do time, a importância de se ter diversidade, pois daí que nasce a pluralidade de ideias. Priorizando as vozes femininas, pois Ana afirma que sente que elas não são tão valorizadas como as masculinas no meio de ambientes de desenvolvimento de tecnologia.

Grupo de foco com homens

Inicialmente, Afonso menciona sua experiência ao criar um dashboard que, aos seus olhos, estava "bem legível, bem legal". A confiança de Afonso na legibilidade e apelo visual do dashboard reflete um entendimento comum entre desenvolvedores com forte embasamento técnico, mas com pouca experiência em *front-end* ou design gráfico. No entanto, ao apresentar o trabalho ao irmão, que tem formação em comunicação e design, Afonso se depara com uma crítica inesperada: o irmão não entendeu o dashboard. a importância de testar e validar os produtos com pessoas de fora do processo de desenvolvimento,

Renato menciona sobre incluir pessoas com outros *backgrounds* em sessões de *design thinking* e *brainstorming* no começo do projeto para entender os problemas e planejar como serão tratados.

6.2

Votação dos indicadores

Pensar a jornada de desenvolvimento de algoritmos com norteadores de desenvolvimento de produtos, segundo os participantes, abriu espaço para que pensassem sob novas perspectivas e incitou um debate sobre questões que não necessariamente estivesse sob o domínio de todos.

Houve sugestões sobre as ordens das perguntas, sobre a abrangência de cada indicador e ao final de cada grupo de foco fizemos uma votação para que os participantes pudessem escolher aqueles indicadores que trouxeram relevância para o debate.

O grupo dos participantes homens, votou conforme a figura 6.3. A maioria declarou que os dois primeiros indicadores, trazendo questões sobre a diversidade, já estão mais latentes na rotina de dados. Os três últimos, que trazem questões de co-desenvolvimento para melhorar a experiência dos usuários, todos afirmaram que estão mais distantes de sua realidade. Talvez por isso, elegeram o terceiro indicador como o mais original e que suscitou uma discussão mais rica, afirmando que pensar em uma experiência adequada pode resolver um pouco das dificuldades apresentadas nos outros indicadores. Ao finalizar a dinâmica, alguns participantes declararam que haviam aprendido bastante com o debate.

O grupo das mulheres, se engajou mais com os direcionadores que falaram concretamente sobre vieses (02) e acessibilidade (05). Tiveram pouca conexão com o direcionador que trazia questões de metodologias ágeis (04). Ao final, depois da votação que ficou conforme a figura 6.4, houve comentários sobre acharem que o debate tinha sido muito significativo, por se sentirem

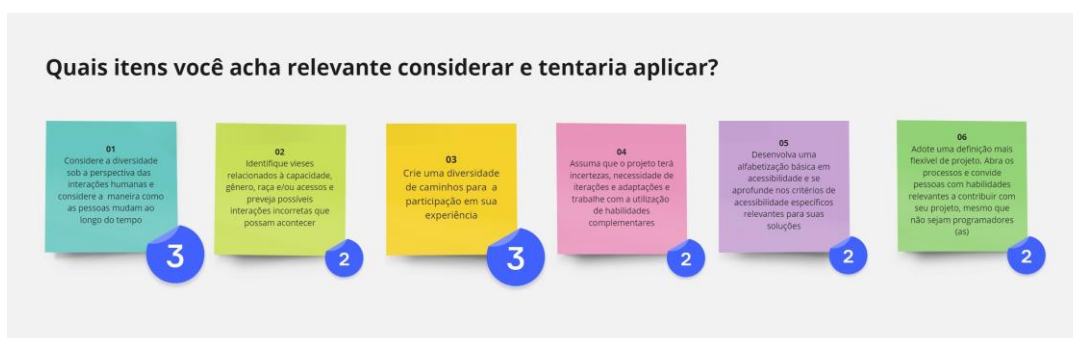


Figura 6.3: Votação dos indicadores - grupo de foco com participantes homens fortalecidos e representados.



Figura 6.4: Votação dos indicadores - grupo de foco com participantes homens

6.3

Observações finais

A principal diferença notada entre os grupos, é que no geral os exemplos trazidos para falar de vieses ou alguma questão de iniquidade, foram feitos em primeira pessoa pelas mulheres. No grupo de homens surgiu um comentário sobre o tratamento diferente dado a mulheres por clientes em ambientes muito técnico, nesse caso, profissionais homens são mais credibilizados. No grupo de mulheres, houve alguns comentários sobre o desconforto vivido por não ter um ambiente com equidade de gênero no trabalho.

Todos os indicadores proporcionaram reflexões acerca de questões de vieses de gênero e situações que demandam mais atenção por fugirem das condições mais gerais. Os primeiros três indicadores geraram mais debates, os dois primeiros por trazerem questões mais explícitas de diversidade, já fizeram com que os participantes inserissem esses temas no debate. O grupo de mulheres trouxe experiências do dia a dia como convivências familiares com filhos, pais ou observação de situações cotidianas que as fazem enxergar especificidades de condições. O grupo de homens trouxe exemplos mais práticos, vividos em alguma experiência profissional. Essa relação pode ser interpretada como um

rebatimento dos papéis institucionalizados de que a mulher fica com trabalhos de cuidado e é mais sensível e homens se dedicam majoritariamente a suas carreiras.

Houve um consenso interessante em relação ao terceiro indicador entre o grupo dos homens. Trazendo a prática de pensar em diferentes maneiras de projetar uma experiência, os participantes elegeram esse item como o que mais os fez pensar de maneira diferente do que eles praticam no dia a dia. O fato de ser uma prática que necessariamente passa por pensar no usuário e nas diferentes formas de interação, segundo os participantes, os pareceu mais original dentre todos os indicadores. Renato destacou como o terceiro ponto servia como uma ponte para resolver problemas de diversidade e acessibilidade apresentados nos outros pontos, sugerindo que compreender e abordar adequadamente um aspecto pode facilitar a solução de outros.

Outro ponto importante, é que ninguém havia falado em situações para pessoas com deficiência ou em questões de acessibilidade durante as entrevistas. No grupo de foco, a temática dos direcionadores as fizeram pensar em exemplos e a discussão trouxe novas realidades de maneira orgânica e muito presente.

De modo geral, todos os participantes se demonstraram atentos aos efeitos que a tecnologia pode trazer e sobretudo, estão conscientes sobre seus papéis nesse sistema.

7

Conclusão

Na linha do tempo dessa pesquisa que se iniciou em 2020, observou-se naquele momento, que havia estudos emergentes sobre vieses reproduzidos pelos algoritmos de IA e seus impactos em geral. A partir de 2022, contudo, já existiam comprovações técnicas de vieses de gênero, raciais, entre outros. Durante a pandemia, todas as interfaces tecnológicas se expandiram significativamente e por conseguinte, a reverberação dessas experiências tomou proporções cotidianas para a sociedade.

O mapeamento sistemático da literatura, já no avançar da pesquisa, revelou que a maioria dos estudos se concentrava no campo tecnológico, onde os vieses eram testados de maneira técnica, com foco nos dados. Notou-se uma lacuna científica no aprofundamento do entendimento da dimensão humana no desenvolvimento dos sistemas. Os especialistas em desenvolvimento de IA eram frequentemente mencionados de forma abrangente e generalista, de acordo com o perfil sociodemográfico. No início da pesquisa, essa era a abordagem predominante. Buscava-se compreender, ainda que de modo generalista, se as intercorrências de experiências de vida refletiam nas decisões implementadas nos projetos. Assim, houve um aprofundamento inicial sobre as relações estruturais das organizações em rede e a mercantilização dos dados, pensando no aproveitamento destes como uma moeda de troca pelo uso de sistemas, e na conscientização do usuário de que, na jornada completa dos dados, ele produz, consome e tem seus dados utilizados para retroalimentar ecossistemas digitais, frequentemente monetizados.

Dessa maneira, para haver a produção de dados, há que se ter acesso, condições de uso da tecnologia, seus sistemas e facilidades. Portanto, no capítulo 2, foram reunidos conceitos para caracterizar os vieses em seus efeitos de fato, como a criação de brechas digitais em observação do paralelo digital, sendo as realidades rebatidas nos meios digitais.

O foco na discussão de iniquidade de gênero surgiu da própria experiência da pesquisadora em ambientes tecnológicos. Essa foi a principal motivação para integrar os três pilares: tecnologia/inteligência artificial, design e gênero com olhar para os impactos decorrentes de ambientes de desenvolvimento predominantemente masculino. Desta forma, o fechamento desse capítulo

se dá no levantamento de aspectos práticos da desigualdade, sobretudo no desenvolvimento e ascensão de carreira, nesse caso, com recorte na carreira acadêmica.

No capítulo 3, são trazidos aspectos técnicos de Inteligência Artificial, para que se crie uma base teórica para o entendimento funcional prático da tecnologia, destrinchando o modelo de desenvolvimento em si. As fases de pré-processamento, o projeto do modelo e questões de aprendizado de dados, demonstram o arcabouço técnico envolvido no desenvolvimento de algoritmos e sistemas com IA. No intuito de esquadrihar a estrutura de algoritmos, são reunidas técnicas de auditoria, que representam um esforço em tangibilizar e explicar as relações entre os códigos e os efeitos em suas aplicações.

Na segunda metade do capítulo 3, os aspectos que envolvem a experiência do usuário e a relação com elementos e princípios de IA, são apresentados. Uma conexão entre aspectos de usabilidade, experiência do usuário e como ela se dá através de pontos de contato com a tecnologia e discussões mais recentes dessa interação são engendrados no conceito de HAI - *Human centered Artificial Intelligence*. A sustentação do conceito de HCD - *Human Centered Design* e seu ecossistema de desenvolvimento, são trazidos na sequência para iluminar todos os agentes na relação humano-máquina envolvidos em situações em que sistemas de IA podem ser experimentados. Nomeia-se, portanto, o conceito de algoritmo centrado no humano, que representa, na forma mais granular, o artefato em questão.

Por fim, dentro da reflexão em que pesa a parte humana na relação tecnologia-homem-sociedade, o conceito de design universal é apresentado, costurando a estrutura teórica com métodos que auxiliam o desenvolvimento de produtos, sistemas e serviços com olhar para a diversidade, acessibilidade e inclusão ampla das necessidades específicas de alguns grupos de usuários.

Ao explicar vieses de gênero, agravamento de iniquidades e experiência do usuário, o exemplo um algoritmo do *facebook* selecionou apenas perfis masculinos para apresentar uma oportunidade de emprego em tecnologia, é uma via diretiva para materializar essa discussão. Logo, já que nesse caso, é possível enxergar com clareza o desenho e os efeitos gerados pelos vieses, esse exemplo foi utilizado como base para o desenvolvimento da metodologia dessa pesquisa.

A pesquisa se volta para o indivíduo decisor, parte do projeto, suas justificativas e modelos mentais. Inicialmente de forma individual e posteriormente de forma coletiva, mas mantendo uma identidade de grupo por gênero, nesse caso, feminino e masculino.

O foco desta pesquisa, portanto, foi analisar os modelos mentais adota-

dos por especialistas em Inteligência Artificial (IA) no processo de criação de algoritmos. Define-se como especialista indivíduos que trabalham profissionalmente no desenvolvimento de algoritmos para sistemas de IA.

Com a discussão centrada na hipótese de que a falta de diversidade poderia limitar a compreensão e a consideração de diferentes necessidades e perspectivas, potencializando a criação de vieses inconscientes nos algoritmos e exacerbando as disparidades digitais. Propõe-se que seja possível tangibilizar marcadores ético-sociais que os especialistas consideram ao desenvolver algoritmos. Portanto, foi elaborada uma lista de indicadores centrados em questões inclusivas e utilizando técnicas de Design, para intensificar o engajamento e o discernimento dos especialistas. Tal abordagem promove maior conscientização durante as tomadas de decisão, especialmente em projetos de grande impacto social, contribuindo para um desenvolvimento mais inclusivo e atento a questões fundamentais.

E como objetivo geral deste estudo, foi delimitada a investigação das diferentes perspectivas de gênero (masculino e feminino), se de fato, influenciam no design de algoritmos. Acredita-se que a falta de diversidade em geral, as limitações na compreensão e consideração de diferentes perspectivas e necessidades pode agravar a geração de vieses inconscientes nos algoritmos e ampliação das brechas digitais.

Para tanto, a metodologia se compõe de entrevistas com método de inquirição contextual, em que um desafio técnico é proposto a cada especialista individualmente, e é pedido que eles narrem todas as suas decisões e argumentações enquanto a mediadora observa seus comportamentos. Foram realizadas 20 entrevistas, 10 com especialistas homens e 10 com especialistas mulheres. A análise dos resultados foi estruturada com base no método de análise das co-ocorrências, a partir de trechos das entrevistas, formando unidades de contexto com a classificação de unidades de registro por categorias: Estruturação Técnica, Aspectos de UX, Princípios de IA e Checklist para desenvolvimento inclusivo. essas categorias representam os elementos de interesse a serem investigados, se havia ou não, a presença e a preocupação em envolver questões mais intrínsecas a essas temáticas na narrativa de cada entrevistado.

No capítulo 5, a sequência da análise segue a organização da Estrutura Técnica, conforme o roteiro das entrevistas. Os trechos das entrevistas mais representativos para cada seção são destacados com a representação visual do grupo de unidades de análise detectados nas falas.

Como parte da investigação, após as entrevistas foram realizados dois grupos de foco, com participantes remanescentes das entrevistas. Um grupo com homens e outro com mulheres. No capítulo 5 é apresentada a técnica e os

direcionadores que foram discutidos pelos grupos, formando o checklist para desenvolvimento inclusivo.

Muito se tem discutido sobre a necessidade de implementar ações afirmativas para mitigar vieses. Nesse contexto, os especialistas frequentemente se deparam com situações que exigem esforços para promover melhorias. Portanto, ações para evitar vieses de gênero, racial e social já haviam emergido no contexto das entrevistas, mesmo sem estímulo. No grupo de foco, quando estimulados, todos trouxeram questões sobre diversidade e acessibilidade, despertando curiosidade sobre a dinâmica de envolvimento com métodos de experiência do usuário.

Ao contactar as especialistas houve um sentimento de grupo despertado, que ficou evidente no grupo de foco. As mulheres trazem as experiências de iniquidade como interlocutoras em primeira pessoa, isso desenha experiências e traz conexões mais potentes.

Quando do início dessa pesquisa se acreditava que poderíamos chegar a um resultado bem mais marcado com apontamento das decisões diferenciadas entre gêneros. No entanto, o crescente discurso de mitigação de desigualdades, surte efeito e a conscientização de que as narrativas devem estar embutidas de consciência social, já está aderida.

O destaque dos resultados vem para o valor que se tira da personificação dos especialistas, o desenho dessas personas confere questões mais profundas para a discussão do que se deve ser melhorado, tanto em questão do próprio ambiente de trabalho, tanto como seus posicionamentos individuais como cidadãos.

Há a sensação que ninguém, enquanto profissional, gostaria de ser responsável por qualquer direcionamento que possa desdobrar em vieses que afetam de diferentes maneiras grupos da sociedade. Portanto, a necessidade de mais trocas entre pensamentos e realidades diversas é colocada pelo grupo investigado, como necessária, e acima de tudo como uma das ações principais para mitigação de vieses. Serem vistos e ouvidos, trouxe uma densidade importante para o olhar mais complexificado da discussão de gênero e iniquidade na tecnologia.

7.0.1

Limitações e oportunidades para futuros trabalhos

As limitações deste estudo, em grande parte, derivam da homogeneidade do grupo testado, o que também é um reflexo das características predominantes nos ambientes de tecnologia. Para facilitar a logística e o engajamento, optou-se por convidar uma equipe mais próxima para participar da pesquisa. Apesar

disso, oferecer vantagens práticas, como maior facilidade de organização e participação ativa, também limita a diversidade de perspectivas que poderiam ser exploradas.

Dada esta limitação, identifica-se uma oportunidade significativa para futuros estudos: a ampliação da diversidade dos grupos envolvidos. Incluir uma variedade mais ampla de gêneros (não binário, fluído, etc.), bem como de backgrounds culturais, poderia enriquecer as discussões e fornecer resultados mais abrangentes e representativos. Promover essa mistura não só alinharia o estudo com práticas inclusivas mais robustas, mas também aumentaria a validade e a aplicabilidade dos resultados, oferecendo uma compreensão mais ampla de como diferentes grupos percebem e são afetados por vieses em ambientes tecnológicos.

Outra oportunidade seria testar de fato o checklist, realizando o desenvolvimento de algoritmos com o apoio dos direcionadores.

Referências bibliográficas

- Agarwal, A., Beygelzimer, A., Dudík, M., Langford, J., & Wallach, H. (2018). A reductions approach to fair classification. *International conference on machine learning*, 60–69.
- Ali, M., Sapiezynski, P., Bogen, M., Korolova, A., Mislove, A., & Rieke, A. (2019). Discrimination through optimization: How Facebook’s Ad delivery can lead to biased outcomes. *Proceedings of the ACM on human-computer interaction*, 3 (CSCW), 1–30.
- Almeida, A. (2015). Design Inclusivo. *Ergotrip Design*, 44–62.
- Arpteg, A., Brinne, B., Crnkovic-Friis, L., & Bosch, J. (2018). Software engineering challenges of deep learning. *2018 44th euromicro conference on software engineering and advanced applications (SEAA)*, 50–59.
- Barbosa, N. M., & Chen, M. (2019). Rehumanized crowdsourcing: A labeling framework addressing bias and ethics in machine learning. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–12.
- Beer, D. (2019). The social power of algorithms. Em *The Social Power of Algorithms* (pp. 1–13). Routledge.
- Bellamy, R. K., Dey, K., Hind, M., Hoffman, S. C., Houde, S., Kannan, K., Lohia, P., Martino, J., Mehta, S., Mojsilovic, A., et al. (2018). AI Fairness 360: An extensible toolkit for detecting, understanding, and mitigating unwanted algorithmic bias. *arXiv preprint arXiv:1810.01943*.
- Bougie, T. (2002). *The impact of new technologies on the quality of life of people with disabilities*. Council of Europe.
- Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New media & society*, 14(7), 1164–1180.
- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Conference on fairness, accountability and transparency*, 77–91.
- Buonocore, T. (2019). Man is to doctor as woman is to nurse: the gender bias of word embeddings. *Towards Data Science*.
- Carpo, M. (2019). Big data and the end of history. *International Journal for Digital Art History: Issue 3, 2018: Digital Space and Architecture*, 3, 21.

- Castells, M. (2011). *The rise of the network society*. John Wiley & Sons.
- Cech, E. A. (2013). The self-expressive edge of occupational sex segregation. *American Journal of Sociology*, 119(3), 747–789.
- Charles, M., & Grusky, D. B. (2005). *Occupational ghettos: The worldwide segregation of women and men* (Vol. 200). Stanford University Press Stanford, CA.
- Cho, W. I., Kim, J., Yang, J., & Kim, N. S. (2021). Towards cross-lingual generalization of translation gender bias. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 449–457.
- Connell, B. R. (s.d.). *Principles of universal design* NC State University. <https://design.ncsu.edu/wp-content/uploads/2022/11/principles-of-universal-design.pdf> (accessed: 03.04.2024).
- Crenshaw, K. (2013). Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. In *Feminist legal theories* (pp. 23–51). Routledge.
- Dev, S., & Phillips, J. (2019). Attenuating bias in word vectors. *The 22nd international conference on artificial intelligence and statistics*, 879–887.
- Dodge, J., Liao, Q. V., Zhang, Y., Bellamy, R. K., & Dugan, C. (2019). Explaining models: an empirical study of how explanations impact fairness judgment. *Proceedings of the 24th international conference on intelligent user interfaces*, 275–285.
- Dubberly, H., & Pangaro, P. (2010). *Introduction to Cybernetics and the Design of Systems*.
- Dutton, W. H., & Reisdorf, B. C. (2019). Cultural divides and digital inequalities: attitudes shaping Internet and social media divides. *Information, communication & society*, 22(1), 18–38.
- Dwork, C., Hardt, M., Pitassi, T., Reingold, O., & Zemel, R. (2012). Fairness through awareness. *Proceedings of the 3rd innovations in theoretical computer science conference*, 214–226.
- England, P. (2005). Gender inequality in labor markets: The role of motherhood and segregation. *Social Politics: International Studies in Gender, State & Society*, 12(2), 264–288.
- Findeli, A. (2001). Rethinking design education for the 21st century: Theoretical, methodological, and ethical discussion. *Design issues*, 17(1), 5–17.
- Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., & Srikumar, M. (2020). Principled artificial intelligence: Mapping consensus in ethical and rights-

- based approaches to principles for AI. *Berkman Klein Center Research Publication*, (2020-1).
- Foulds, J. R., Islam, R., Keya, K. N., & Pan, S. (2020). An intersectional definition of fairness. *2020 IEEE 36th International Conference on Data Engineering (ICDE)*, 1918–1921.
- Gajane, P., & Pechenizkiy, M. (2017). On formalizing fairness in prediction with machine learning. *arXiv preprint arXiv:1710.03184*.
- Galison, P. (1994). The ontology of the enemy: Norbert Wiener and the cybernetic vision. *Critical inquiry*, 21 (1), 228–266.
- Gamboa, S. S. (1997). Revolução informacional: pontos de vista para o debate sobre a sociedade da informação. *Transinformação*, 9(1).
- Ganhör, C., Penz, D., Rekabsaz, N., Lesota, O., & Schedl, M. (2022). Unlearning protected user attributes in recommendations with adversarial training. *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2142–2147.
- Gillespie, T. (2014). The relevance of algorithms. *Media technologies: Essays on communication, materiality, and society*, 167(2014), 167.
- Gillespie, T., Boczkowski, P. J., & Foot, K. A. (2014). *Media technologies: Essays on communication, materiality, and society*. MIT Press.
- Gitelman, L. (2006). Always already new. *Media, history, and the data of culture*, 7.
- Gordano, C. (2009). Construyendo sentido sobre internet en el espacio de la diáspora: mujeres latinas inmigrantes en Granada.
- Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G.-Z. (2019). XAI—Explainable artificial intelligence. *Science robotics*, 4 (37), eaay7120.
- Hassenzahl, M. (2008). User experience (UX) towards an experiential perspective on product quality. *Proceedings of the 20th Conference on l'Interaction Homme-Machine*, 11–15.
- Holtzblatt, K., Wendell, J. B., & Wood, S. (2004). *Rapid contextual design: a how-to guide to key techniques for user-centered design*. Elsevier.
- Invisível, C. (2020). *Aos nossos amigos Crise e Insurreição*. n-1 edições. Jurno, A., et al. (2018). Algoritmos como um devir: uma entrevista com Taina Bucher. *Parágrafo*, 6 (1), 165–170.
- Kalinowski, M., Escovedo, T., Villamizar, H., & Lopes, H. (2023). *Engenharia de Software para Ciência de Dados: Um guia de boas práticas com ênfase na construção de sistemas de Machine Learning em Python*. Casa do Código.

- Kim, B., & Doshi-Velez, F. (2021). Machine learning techniques for accountability. *AI Magazine*, 42 (1), 47–52.
- Kitzinger, J., & Barbour, R. (1999). *Developing focus group research: politics, theory and practice*. Sage.
- Krippendorff*, K. (2004). Intrinsic motivation and human-centred design. *Theoretical Issues in Ergonomics Science*, 5(1), 43–72.
- Kumar, N., & Tanwar, S. (2020). Multimedia Big Data Computing for IoT Applications Concepts, Paradigms and Solutions.
- Kusner, M. J., Loftus, J., Russell, C., & Silva, R. (2017). Counterfactual fairness. *Advances in neural information processing systems*, 30.
- Langlois, G. (2013). Participatory culture and the new governance of communication: The paradox of participatory media. *Television & New Media*, 14 (2), 91–105.
- Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard university press.
- LÉVY, P. (2003). A Inteligência Coletiva: Por uma Antropologia do ciberespaço. 5ª edição. Tradução de Luiz Paulo Rouanet. São Paulo: Loyola.
- Li, F., & Etchemendy, J. (2018). A common goal for the brightest minds from Stanford and beyond: Putting humanity at the center of AI.
- Lutz, C. (2019). Digital inequalities in the age of artificial intelligence and big data. *Human Behavior and Emerging Technologies*, 1 (2), 141–148.
- Mace, R. L. (1996). Accessible environments: Toward universal design. *AUED*, 1996.
- Maguire, M. (2001). Methods to support human-centred design. *International journal of human-computer studies*, 55 (4), 587–634.
- Mayer-Schonberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.
- Melchiorre, A. B., Rekabsaz, N., Parada-Cabaleiro, E., Brandl, S., Lesota, O., & Schedl, M. (2021). Investigating gender fairness of recommendation algorithms in the music domain. *Information Processing & Management*, 58 (5), 102666.
- Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the national academy of sciences*, 109 (41), 16474–16479.
- Natansohn, G. (2014). Mulheres e uso das TIC, algumas reflexões metodológicas. *En Anais do Intercom-Sociedade Brasileira De Estudos Interdisciplinares Da Comunicação. Foz do Iguaçu: Intercom. Recuperado de:*

- <http://gigaufba.net/wp-content/uploads/2014/09/Mulheres-USO-TIC-Intercom2014Natansohn2014.pdf>.
- O'neil, C. (2017). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
- Pasquale, F. (2011). Restoring transparency to automated authority. *J. on Telecomm. & High Tech. L.*, 9, 235.
- Petzinger Jr, T. (1996). *Hard landing: The epic contest for power and profits that plunged the airlines into chaos*. Currency.
- Redmond, K. C., & Smith, T. M. (2000). *From whirlwind to MITRE: The R&D story of the SAGE air defense computer*. MIT Press.
- Rezende, S. O. (2003). *Sistemas inteligentes: fundamentos e aplicações*. Editora Manole Ltda.
- Russell, S., Norvig, P., & Intelligence, A. (1995). A modern approach. *Artificial Intelligence. Prentice-Hall, Egnlewood Cliffs*, 25(27), 79–80.
- Sanders, E. B.-N., & Stappers, P. J. (2008). Co-creation and the new landscapes of design. *Co-design*, 4(1), 5–18.
- Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014). Auditing algorithms: Research methods for detecting discrimination on internet platforms. *Data and discrimination: converting critical concerns into productive inquiry*, 22(2014), 4349–4357.
- Silva, T. (2019). Teoria Racial Crítica e Comunicação Digital: conexões contra a dupla opacidade. *Anais do 42º Congresso Brasileiro de Ciências da Comunicação. Sociedade Brasileira de Estudos Interdisciplinares da Comunicação. Belém, Pará*.
- Simon, H. A. (1996). *The sciences of the artificial*. MIT press.
- Stephanidis, C. (2001). User interfaces for all: New perspectives into human-computer interaction. *User interfaces for all-concepts, methods, and tools*, 1 (1), 3–17.
- Striphas, T. (2010). How to have culture in an algorithmic age. *The Late Age of Print Blog*.
- Sun, T., Gaut, A., Tang, S., Huang, Y., ElSherief, M., Zhao, J., Mirza, D., Belding, E., Chang, K.-W., & Wang, W. Y. (2019). Mitigating gender bias in natural language processing: Literature review. *arXiv preprint arXiv:1906.08976*.
- Tessier, M. (2021). *Digital Suffragists: Women, the Web, and the Future of Democracy*. MIT Press.
- Verganti, R., Vendraminelli, L., & Iansiti, M. (2020). Innovation and design in the age of artificial intelligence. *Journal of Product Innovation Management*, 37(3), 212–227.

- Wajcman, J. (2006). *El tecnofeminismo*. Universitat de València.
- Wang, C., Wang, K., Bian, A., Islam, R., Keya, K. N., Foulds, J., & Pan, S. (2022). Do humans prefer debiased AI algorithms? A case study in career recommendation. *27th International Conference on Intelligent User Interfaces*, 134–147.
- Weisshaar, K. (2017). Publish and perish? An assessment of gender gaps in promotion to tenure in academia. *Social Forces*, 96(2), 529–560.
- Wiener, N. (2019). *Cybernetics or Control and Communication in the Animal and the Machine*. MIT press.
- Witness, G. (s.d.). *New evidence of Facebook's sexist algorithm*. <https://www.globalwitness.org/en/campaigns/digital-threats/new-evidence-of-facebooks-sexist-algorithm/> (accessed: 07.09.2023).
- Wolton, D. (1999). *Pensar a comunicação*. Difel - Difusão Editorial, S.A.
- Xu, W. (2019). Toward human-centered AI: a perspective from human-computer interaction. *interactions*, 26(4), 42–46.
- Young, I. (2008). *Mental models: aligning design strategy with human behavior*. Rosenfeld Media.
- Zahm, M. (1997). Advances in Human-Computer Interaction, Volume 5 Edited by Jakob Nielsen 1995, 258 pages, Norwood, NJ: Ablex Publishing ISBN 0-56750-196-6. *Ergonomics in Design*, 5(3), 35–35.
- Zuboff, S. (2019). The age of surveillance: the fight for a human future at the new frontier of power. *New York: Public Affairs*.

A

Apêndice: Termo de Consentimento Livre e Esclarecido

LEUI| Laboratório de Ergodesign e Usabilidade de Interfaces da PUC-Rio

Título da Pesquisa: Algoritmo centrado no humano: os caminhos para decisões de projeto

Nome da Pesquisadora Responsável: Ma. Larissa Pereira Marques Cruz

Professora Orientadora: Profa. Dra. Manuela Quaresma

Você está sendo convidado/a a participar da pesquisa intitulada Algoritmos centrado no humano: os caminhos para decisões de projeto, sob a responsabilidade da pesquisadora Ma. Larissa Cruz aluna de doutorado do curso de Doutorado em Design da Pontifícia Universidade Católica do Rio de Janeiro (PUC- Rio), sob a orientação da Profa. Dra. Manuela Quaresma.

Justificativa

Nossa pesquisa investiga se a diversidade dos profissionais de tecnologia (especialistas em desenvolvimento de IA) traz impacto no projeto de modelagem dos algoritmos e se, por consequência, pode levar a reprodução de vieses algoritmos inconscientes ampliando as brechas digitais e a desigualdade de gênero em STEAM.

Objetivo

O objetivo dessa entrevista é evidenciar as decisões de projeto empregadas pelos (as) profissionais que possuem vivência e experiência em modelagem de algoritmos. Tencionamos investigar quais recursos, métodos, escolhas, referências apoiam as decisões dos profissionais e de maneira mais sensível, pretende-se correlacionar suas ações a seus contextos sociais demográficos.

Metodologia

A entrevista será realizada em áudio e vídeo. Você receberá um link para uma plataforma de videoconferência na qual transcorrerá a entrevista. Muito embora sua entrevista vá ser gravada para permitir posterior análise, nosso interesse reside apenas no conteúdo de suas respostas. As partes do seu depoimento que forem mais pertinentes ao nosso interesse acadêmico serão

transcritas. Os dados serão confidenciais, restritos apenas aos pesquisadores. Estimamos que a entrevista durará cerca de 40 minutos e abordaremos o seu comportamento como profissional desenvolvedor/a de modelos algorítmicos.

Participantes da pesquisa

Nosso público-alvo é composto por profissionais com experiência em desenvolvimento de modelos algorítmicos. Você foi selecionado pelo método bola de neve em que profissionais da área de tecnologia, através de contato prévio, indicaram seu nome como apto/a para participar dessa entrevista

Riscos

Esta pesquisa tem como riscos: possíveis desconfortos ou constrangimentos em compartilhar suas opiniões. Mas, para diminuir a chance desses riscos acontecerem, reforçamos que suas respostas serão totalmente anônimas, e não será possível identificar sua identidade no conteúdo gerado, durante ou após a realização do estudo. O procedimento do qual participa visa somente criação e discussão de soluções e, logo, você não será testado(a) ou julgado(a).

Benefícios

Sua participação é muito importante para a compreensão do comportamento profissional desenvolvedor/a de modelos algorítmicos, o que poderá lhe beneficiar futuramente de forma indireta.

Sigilo e Privacidade

Para proteger o sigilo de sua identidade, seu nome não aparecerá em nenhuma publicação. Você receberá um pseudônimo (um nome falso) que será usado em vez de seu nome. Todo o material de áudio e vídeo será tratado como confidencial e restrito para fins acadêmicos. Além disso, como já exposto, os registros de áudio e vídeo serão utilizados apenas para permitir a análise do conteúdo de suas respostas e não serão divulgados de nenhuma forma.

Autorização para uso de declarações

O material que constitui o corpo de dados coletados (imagens e áudio) não será divulgado. Você autoriza o uso de suas declarações e imagens para finalidades acadêmicas – artigos acadêmicos, aulas, papers, sites, apresentações em simpósios ou congressos científicos relacionados ao tema?

() autorizo a divulgação de minhas declarações em formato de texto, sem uso da minha voz.

☐ autorizo a divulgação de minha imagem apagando a face

☐ não autorizo

Despesas e Ressarcimento

Para participar deste estudo, você não terá nenhum custo e nem receberá qualquer vantagem financeira.

Assistencia e acompanhamento

Os resultados da pesquisa estarão à sua disposição quando finalizada, bastando solicitar uma cópia por e-mail ao pesquisador. Seu nome ou o material que indique sua participação não será liberado sem a sua permissão. Você não será identificado (a) em nenhuma publicação que possa resultar. Caso eventualmente a pesquisa seja interrompida, seus dados não serão publicizados de nenhuma forma.

Participação voluntária e direito de desistência

Sua participação nesta entrevista é voluntária. Você não tem nenhuma obrigação de participar. Se você necessitar de uma pausa a qualquer momento durante a entrevista, por favor, avise. Você tem o direito de mudar de ideia e sair da entrevista a qualquer momento, sem apresentar motivos e sem qualquer penalização. Qualquer nova informação que possa fazê-lo(a) mudar de ideia sobre estar na pesquisa será fornecida a você. Este documento será emitido em duas vias, sendo uma para as pesquisadoras e outra para você quando assinadas. Para confirmar a sua autorização via e-mail, deverá responder para larissapmc@gmail.com, com a seguinte informação expressa: estou ciente do termo de consentimento e aceito participar. Você não renuncia a qualquer de seus direitos legais ao assinar ou concordar com este termo de consentimento.

Acesso à pesquisadora em caso de dúvidas

Em qualquer etapa da pesquisa, você pode contatar a pesquisadora responsável para esclarecimento de dúvidas pelo e-mail larissapmc@gmail.com ou pelo telefone (21) 99191-9481 ou com a professora orientadora através do e-mail mquaresma@puc-rio.br ou pelo telefone (21) 3527-1005.

Câmara de Ética em Pesquisa da PUC-Rio

Em caso de dúvidas relacionadas à ética em pesquisas, acionar o órgão responsável na Rua Marquês de São Vicente, No 225, Edifício Kennedy, 2º andar,

Gávea - 22453-900. Rio de Janeiro, RJ. Telefone: (21) 3527-1618

Resoluções 466/12 e 510/16 do CNS

Esta pesquisa atende todas as especificações da Resolução 466, de 12 de dezembro de 2012, que aprova as diretrizes e normas regulamentadoras de pesquisas envolvendo seres humanos, e da Resolução 510, de 07 de abril de 2016, que dispõe sobre normas aplicáveis à pesquisa em Ciências Humanas e Sociais.

Afirmo que li, compreendi e concordo em participar da pesquisa e que me foi dada a oportunidade de esclarecer minhas dúvidas.

Data: _____, _____ de _____ de _____.

Assinatura do Participante

Assinatura da pesquisadora responsável:

Larissa Pereira Marques Cruz

Assinatura da professora orientadora:

Manuela Quaresma

Departamento de Artes e Design

Programa de Pós-graduação em Design da PUC-Rio

LEUI | Laboratório de Ergodesign e Usabilidade de Interfaces

Rua Marquês de São Vicente, 225, sala 711F – Edifício Frings

Gávea, Rio de Janeiro, RJ. CEP: 22453-900. Telefone: (21) 3527-1005

Câmara de Ética em Pesquisa da PUC-Rio

Rua Marquês de São Vicente, 225 – Edifício Kennedy, 2º andar. Gávea, Rio de Janeiro, RJ.

CEP: 22453-900. Telefone: (21) 3527-1618. E-mail: alef@puc-rio.br.

A Câmara tem por atribuição analisar do ponto de vista ético os projetos de pesquisa dos professores, pesquisadores e discentes da Universidade, quando solicitada.

B

Apêndice: Entrevistas

B.1

Afonso

Afonso: Mesmo assim, botar em inglês tudo que você fala em português, fica uma zona.

Entrevistadora: Sim, exatamente. Já aconteceu isso, eu esqueci de trocar, aí ficou uma beleza. Deixa eu só colocar aqui a linguagem.

Afonso: E de público-alvo, mas as pessoas que você está entrevistando estão no mesmo contexto assim? São pessoas de áreas diferentes, idades, formações?

Entrevistadora: É, o que que eu tô... gente, onde que eu... não tô achando onde é o negócio aqui.

Afonso: Eu posso colocar aqui pra gravar.

Entrevistadora: Não, eu botei, agora eu tô tentando trocar a linguagem e eu não tô lembrando onde é o...

Afonso: Deixa eu ver... um momento... Em línguas diferentes... Achei! Vou trocar aqui, tá?

Entrevistadora: Tá.

Afonso: Português Brasil... Teste...

Entrevistadora: Boa! Foi. Muito bom, obrigada.

Afonso: De nada.

Entrevistadora: O que é que... sobre as pessoas, né? Eu vou... a simulação, né, da solução aqui que eu vou pedir pra você fazer um exercício de pensar o que você faria, né, as decisões, é sobre... pra fazer um algoritmo de recomendação. Então, assim... Eu pensei que eu deveria buscar... trabalhar, assim, só com algoritmo de recomendação.

Travou?

Afonso: Travou, voltou agora.

Entrevistadora: Voltou. A princípio eu pensei que só deveria pegar pessoas que trabalhassem com recomendação, né, algoritmo de recomendação.

Mas eu comecei a fazer com pessoas próximas que trabalham com algo... desenvolvem algoritmos que não necessariamente precisam ser de recomenda-

ção, mas que tem o racional pra fazer um... pra desenvolver uma solução que use algoritmo. Ou que tenha alguma vivência com um modelo, sabe, que tenha essa... é o racional mesmo. E aí tá sendo interessante porque tá sendo bom, assim, porque aí as pessoas não têm muito vício, sabe? E aí a ideia é, tipo, você liberar aí seus pensamentos.

Você recebeu essa demanda, né, como se você tivesse recebido essa demanda e é tipo e agora, o que eu faço? E aí vai pensando e me dizendo. A gente tem no design, né, que eu tô fazendo doutorado no design especificamente. A gente chama... tem esse método de Directed Storytelling, que é você ir dizendo o que você faria.

E por isso que eu falei, se você quiser compartilhar também a sua tela em algum momento, fica à vontade, mas não precisa também, não é necessário, não é obrigatório. E aí você vai me dizendo mesmo, se eu faria isso. É ir narrando o seu pensamento, tá bom? Vou compartilhar aqui.

E aí a primeira coisa, até para eu entender melhor, assim, sobre você. Eu vou pedir para você contextualizar a experiência, a sua experiência com algoritmos e mais especificamente também, assim, a sua experiência profissional, o que você trabalha hoje, para a gente ter uma contextualização.

Afonso: Ótimo, Larissa.

Então, primeiramente eu sou formado em Engenharia Eletrônica e de Computação pela UFRJ. E desde a minha formação eu tenho gostado mais da parte de Computação. Então, vi a questão da Eletrônica e da Computação no curso, na graduação, só que me interessei mais na Computação.

Foi onde eu comecei a fazer meu estágio, até chegar aqui hoje na Deloitte. E no percurso, eu fui bastante entre áreas inteiras da Programação. Comecei com o Banco de Dados, fui para a área de Mobile, desenvolvimento de aplicativo.

E depois eu comecei muito na parte de desenvolvimento de back-end, de fazer algumas APIs, alguns serviços. E eu tenho uma base forte nessa questão de estruturação de código, performance, base de programação como um todo. E foi ali que eu aprendi bastante coisa sobre algoritmo, de como você conseguir organizar seus pensamentos para conseguir resolver um problema.

Não sair fazendo como muito programador faz. Então, existe até esse conceito de programador e desenvolvedor. O programador chega e vai sair programando.

O desenvolvedor já pensa em como pode fazer, faz uma análise crítica do que vai acontecer se você implementar daquela forma. E depois dessa questão, depois de alguns anos nessa área, eu tenho uns cinco anos de mercado, eu comecei mais na parte de engenharia de Dados. Eu tenho um pouco de

formação na parte de Ciência de Dados, mas não é hoje o que eu faço.

Eu nunca estou no mercado com Ciência de Dados, mas eu tenho um MBA em Ciência de Dados. E eu comecei a me interessar tanto na questão do modelo, desenvolvimento de software e acabei indo para a área de engenharia de Dados. Então, aqui na empresa, eu estou já uns dois anos e um pouquinho na área de engenharia de Dados, onde eu penso mais na arquitetura de como resolver um problema.

Então, faço integração de sistemas. Se aquilo que estou fazendo é escalável ou não, se vai atender à solução ou não, eu estou com foco muito nessa questão. De pegar o dado ali da fonte, preparar o dado e disponibilizar isso para o cientista ou para outras pessoas que trabalharam com os dados.

Entrevistadora: Entendi.

Afonso: Então, sobre a questão de algoritmo de recomendação ou outro algoritmo de Ciência de Dados, nunca trabalhei assim no mercado, mas conheço algumas coisinhas bem soltas. Não é muito a minha área, mas já corri um pouquinho atrás sobre.

Entrevistadora: Muito bom. Tá legal. Eu vou te passar o nosso objetivo aqui, te contextualizando sobre o cenário.

A minha ideia é o seguinte. Vamos dizer que você trabalhe numa empresa que faça algoritmos, que faça algoritmos para recomendação. Enfim, nesse caso específico, uma empresa de RH, desculpa, uma empresa de tecnologia, te contrata, contrata a sua empresa para fazer a divulgação do anúncio de uma vaga de trabalho.

Então, tem uma vaga aberta, eles te passam somente esses requisitos, que é uma vaga de tecnologia para profissional pleno, com skills de desenvolvedor, full-stack, ou desenvolvedora, com a formação em engenharia de software, tecnologia, sistemas, informação na parte de ciências mesmo, tecnologia, uma experiência de cinco anos e regime remoto. Aí, assim, só para começar a pensar aí, se for desenhar a solução, imagina você num Facebook da vida, aqueles anúncios que ficam aparecendo, seria você programar esses anúncios com a vaga para eles aparecerem para pessoas que tenham um perfil que pode ser que se interesse por essa vaga. Daí, assim, não precisa, você não precisa fechar em uma plataforma, como a gente está falando de um cenário fictício, mas te ajudar a pensar num LinkedIn, num Facebook, seria uma coisa por aí.

Então, na primeira etapa aqui, a gente quer saber como você faria essas definições de requisitos a partir dos requisitos que te passaram, que são aqueles do slide anterior, e quais seriam os dados utilizados, como você faria essa seleção de dados a partir daqueles tópicos que te passaram?

Afonso: Um momento, deixa eu só separar a pergunta aqui, eu vou pedir

para você botar no slide anterior, e eu vou ficar vendo a pergunta aqui para abordar todos os temas. Espera aí, deixa eu jogar isso aqui para cá. Tá, bom, vamos lá.

Quando coloca uma empresa de tecnologia, eu acho que isso já define muita coisa hoje, tá? Porque, assim, dependendo da área de datação da empresa, existem certas plataformas, um certo público que você quer atingir, existem certas plataformas para tal. Apesar de ter ferramentas hoje no mercado, como o LinkedIn, acho que é o mais famoso que tem, atende todas as áreas possíveis, a parte de design, qualquer área atende, mas eu entendo que certas áreas do mercado ainda não têm essa mente de utilizar o LinkedIn, por exemplo. E eu vivenciei isso com minha esposa aqui, no interior do Rio, então é muito mais currículo entregar de modo físico na porta da empresa, mas como a gente está falando de tecnologia e porte médio, se fosse porte pequeno, até poderia considerar esse caso, a empresa mais tradicional de interior, mas com a porte médio, já presumo que vai ter pessoas ali que estão mais atentos ao mercado, o que o pessoal está utilizando e tudo mais.

Você comentou de Facebook, mas no meu caso, o que passa na minha cabeça seria através do LinkedIn. E a empresa de tecnologia tem muito forte a parte de recrutamento com o que eles chamam de Headhunter. Novamente, se não fosse pela tecnologia, isso não ia acontecer, ter um Headhunter não enxerga muito nas outras áreas, mas a tecnologia é muito forte.

De ter alguém nessa empresa que está sendo contratada, de utilizar o plataforma do LinkedIn para procurar aquele profissional com esses requisitos. Então, a primeira plataforma sendo atendida vai ser o LinkedIn, por exemplo. Também tem o JobConvo, o Guppy, mas eu entendo que a porta de entrada ali de empresa de porte médio para cima seria mais o LinkedIn.

Existem outras plataformas, mas eu vejo mais como um porte médio para pequeno. Tem o InfoJobs, tem o Catto também, que é até pago, se não me engano, mas na nossa área, sim, em empresas maiores, eu vejo utilizando muito mais o LinkedIn. Então, seria a abordagem que eu iniciaria através do LinkedIn.

Bom, sobre a vaga, as informações que foram passadas pela empresa, eu não sei se poderia, se a empresa que está sendo contratada, poderia, como eu posso falar, comentar a respeito, se é flexível ou não esses requisitos. Porque, por exemplo, pleno, com cinco anos de formação, talvez teria uma dificuldade de encontrar, tá? Talvez seria interessante entender por que são cinco anos de formação para um pleno. Então, por exemplo, no meu caso, eu tenho cinco anos e sou sênior, dois indo para o três.

Outras pessoas também que estão nessa faixa já são sênior, né? Então,

encontrar uma pessoa que está há cinco anos de experiência e ainda é pleno, seria um pouco difícil. Mas aí buscar entender melhor o porquê desse mínimo de experiência. Se é obrigatório mesmo, ou é só uma base? Ah, se você fizer nesses cinco anos aqui, se for cinco anos, se for quatro, aí você analisa a experiência da pessoa e tudo mais, do candidato.

Bom, muito importante também saber, não somente a skill, tá? Desenvolver do Full-stack, mas relacionado a quê? A que tipo de produto? Acho que isso pesa muito, tá? Pode ser uma API, pode ser um outro tipo de serviço, aplicação web, literalmente. Então, já ter alguém que tem que saber mexer com front, com design. A parte do design já é a questão da experiência do usuário, saber como que vai apresentar aqueles componentes ali da melhor forma para ficar legível, entendível pelo usuário.

Então, isso já puxa uma série de outros skills que se não for uma aplicação web e que precisa também desenvolver do Full-stack, tem uma outra série de requisitos técnicos, de skills técnicas. Então, seria importante ter essa questão. Formação hoje, né? Eu vejo também que não vejo como muito obrigatória, sim.

Aí também entrar em contato com a empresa vai ser realmente necessário. Pode ser que o produto que ela vende para um cliente ou essa empresa também fornece algum serviço para o cliente, para o outro cliente, né? E existe essa formação dos profissionais, como a gente passa aqui na empresa também. Tem certas empresas que exigem formação na área.

Então, isso já explicaria o motivo desse requisito. Porque na nossa área não é tão necessário assim. Eu vejo bem importante para o crescimento da pessoa.

Mas isso também talvez restringiria algumas pessoas que seriam tão bem capacitadas quanto. E que com esse filtro você nem olharia, né? Bom. Aliado a essa questão, por exemplo, região remota também.

Hoje está muito difundido aqui entre esse setor. Muitas pessoas não consideram outra possibilidade de regime. Mas o lado interessante disso também que através da região remota você consegue alcançar também pessoas de outros estados que presencial seria inviável, né? Ainda mais por porte médio.

Não ter a capacidade financeira de trazer uma pessoa de outro estado bancar para trabalhar na sua empresa. Não faz sentido. Bom, aí com essas informações alinhadas no LinkedIn.

Eu utilizaria algumas palavras-chave, né? Para pesquisar, de algumas tecnologias, então entendendo essa vaga para que tipo de projeto trabalho. Não necessariamente entrar nos detalhes, isso pode ser um... Não seja necessário. Mas para entender quais skills são necessários.

Porque Full-stack é muito, é muito amplo.

Entrevistador: Aberto, né?

Afonso: Exato. Porque tem a parte reética, Node.js, uma questão. C Sharp, Python é o outro. Então são skills bem diferentes. E precisa entender isso para saber como pesquisar isso no LinkedIn.

Aí tendo essas informações, se for possível, né? Aí eu coloquei no LinkedIn, por exemplo, algumas palavras-chave nesse sentido, de algumas tecnologias, linguagens, ferramentas, para aparecer as pessoas ali que tem isso na descrição, em alguma parte do perfil do LinkedIn. Mas na experiência, aí seria mais um trabalho de cinco anos. Número de... de anos de experiência, não tem, eu acho, um campo assim muito bem definido no LinkedIn.

Teria que analisar a vida profissional da pessoa. Mas colocando palavras-chave como pleno, sênior, que cinco anos não vai ser júnior ainda. Na teoria não deveria ser.

Aí coloquei pleno, sênior na pesquisa também para conseguir filtrar pessoas que têm mais tempo de mercado. Para conseguir filtrar e jogar melhor.

Entrevistadora: Sim.

Afonso: Aí, sobre que tipo de dados você utilizaria para alimentar algoritmos? Como se os dados foram coletados e processados?

Entrevistadora: Como seriam, né? Assim, por exemplo, você me falou algumas questões, eu achei até interessante, assim, até te perguntar. Nesse caso, quando você receber esses requisitos, você ia querer falar com os contratantes, com os clientes, para entender melhor, assim, está muito vago isso aqui, vamos entender, vamos destrinchar. Aí, a partir do momento, vamos dizer que eles te deram todas as definições, tá? Destrincharam ali as skills, colocaram o foco, né? Colocaram, é, não, cinco anos é muita coisa, é só porque a gente não quer uma pessoa muito júnior e tal, mas assim, dá para fazer uma margem aí de experiência.

Te detalharam, tá? O que eles queriam. A partir daí, quando a gente fala de dados, é assim, quais seriam os dados, e aí não precisa se atrelar tanto à solução técnica, tá? Mas quais seriam os dados que te trariam uma informação para conseguir entender melhor, selecionar as pessoas que estão sendo desenhadas aqui, com esse perfil, né? Vamos dizer que eles colocaram realmente, destrincharam todos os perfis e tal, quais seriam os dados? Você até chegou a falar um pouco ali, falou assim, ah, essa parte de experiência não é uma questão direta, num LinkedIn que eu vou olhar, vou olhar só o começo e o fim, não sei, mas como você faria?

Afonso: Tá, um ponto importante também, Larissa, que eu esqueci de mencionar, que eu tenho um lado muito mais técnico da coisa, assim, não tenho

muita área de gestão, assim, mas eu entendo que, tá, talvez não, a empresa contratante também tem um intervalo que ela consegue trabalhar em questão de remuneração, que não está explícito aí, né? Então, não adianta tanto a equipe que foi contratar, a empresa que foi contratada, marcar uma entrevista com as pessoas, sendo que não sabe muito bem a base que a contratante trabalha com a remuneração.

Entrevistadora: Sim, quando eu falei... Desculpa.

Afonso: Pode falar.

Entrevistadora: Não, quando eu falei assim, não se preocupou tanto com a parte técnica, assim, não é, é justamente a parte técnica que eu quero saber, mas eu estou querendo dizer assim, você não precisa, porque, assim, eu imagino que você não tenha conhecimento de como é o LinkedIn, a organização dos dados e essa questão, não tem problema, entendeu? Mas seria assim, realmente, como você montaria, um modelo para rastrear essas pessoas que podem ter um match com esses requisitos? Aí algumas vão ter mais, outras menos, mas você também não teria que entrevistar essas pessoas, é como se você fosse passar uma peneira e aí você devolve a peneira cheia para o contratante, para o cliente.

Afonso: Tá, bom, vamos lá.

É, sobre a questão da experiência, eu olharia a vida profissional da pessoa na plataforma, e isso também com áreas relacionadas com a vaga, então, se a pessoa migrou de área, eu não consideraria tanto isso, apesar de ter algo ali que possa ser aproveitado, eu não levaria muito em consideração. Então, das empresas que a pessoa passou, está na faixa acadêmica, né? Mas, principalmente, na faixa da empresa, que é onde a experiência, literalmente, é de mercado, eu olharia, tá, qual dessas oportunidades que a pessoa passou são relacionadas com a vaga em si, na empresa contratante? E ali na plataforma também aparece o tempo de dedicação da pessoa em cada oportunidade que ela passou na vida, e com isso eu sei se ela tem aquela quantidade de tempo de experiência que a empresa pede. Mas, novamente, em áreas relacionadas, então, se, por exemplo, a Full-stack e a pessoa tem quatro anos de mobile, a pessoa aprendeu muita coisa ali de programação que vai ser útil também, só que não bate muito com Full-stack, é uma outra área interna de tecnologia, então eu ficaria um pouco um pé atrás.

Só se eu não encontrasse outros candidatos, aí eu já começaria a ficar mais flexível. Mas, relacionado a Full-stack, então, eu estou falando de front, back-end, pessoa que fez API, fez algum microserviço, aí eu já apontaria esses anos nessas oportunidades que a pessoa passou, para checar cinco anos nesses cinco anos. Mas, novamente, sabendo se pode ser flexível ou não.

Sobre a formação, normalmente, se for um item obrigatório, eu comecei

com o contratante, é um item obrigatório, então eu procuraria naquela parte acadêmica, na plataforma, não só ali que a gente tem, nos outros demais também tem, nessa área, no grupo das pessoas, se tem algo ligado à tecnologia, então, por exemplo, análise de sistemas, ciência da computação, engenharia da computação, sistemas de informação, então, todas essas áreas, esses cursos que são focados em programação do mesmo software, eu levaria em consideração. Sobre as skills, eu já mencionei, se a pessoa fez alguma API, alguma biblioteca, algum site, isso tudo está relacionado à parte do desenvolvimento do Full-stack, então, são skills que o pâncreas tem esse conhecimento, seria interessante se ele respeitar os outros requisitos, de entender melhor o que a pessoa fez. Até para saber se enquadra como um pleno.

E sobre a vaga para pleno, tem a questão dos anos de experiência, que de certo modo diz alguma coisa assim, já passei por alguns casos que a pessoa tem muitos anos de experiência, mas chega lá na hora, parece que não é mesmo a pessoa, isso acontece, mas não tem como muito bem evitar isso, mas no pleno, olharia muito como a pessoa abordou aquela oportunidade, tem pessoas que não conseguem aproveitar muito aquele campo de descrição do que atuou naquela empresa, isso acaba dificultando, mas olharia bastante aquele campo de descrição, então, se a pessoa liderou, pleno, mas liderou algum projeto ou foi responsável por fazer, por desenvolver algum projeto dentro da empresa, isso é mais característico de pleno. Se a pessoa também meio que assumiu um projeto mesmo que já existia, mas está de frente ali com o cliente, alinhando algumas coisas, isso também eu vejo como algo positivo para aquele candidato ser entrevistado, mas se a pessoa só deu manutenção em algum projeto, não abordou nada com comunicação com o cliente ou que fez um, ou que enfrentou um problema ou entendeu que entrou para atuar num certo problema que é complexo, aí eu ficaria, será que essa pessoa seria plena? Mas analisaria também os anos de experiência, as outras oportunidades que a pessoa descreveu para saber se enquadra ou não. Então, são esses tipos de dados que eu coletaria para saber se a pessoa encaixa ou não nessa vaga.

Entrevistadora: Tá.

Afonso: E sobre o regime remoto, primeiro ponto, se eu vejo que a pessoa em algum momento atuou como remoto, talvez a pessoa coloque ali na descrição, remotamente, home office, em algum lugar, uma pessoa que já começou, já trabalhou remoto, a tendência é ela continuar remoto. Assim eu entendo, tá? Então, se aparecer como remoto, a tendência é dela continuar como remoto.

Então, atenderia essa vaga a partir do regime remoto. Assim como... Esse é um requisito da contratante que é até bem aceitável para as pessoas. Então,

seria oposto para o presencial.

Aí seria muito mais difícil de lidar. A pessoa já trabalhou remoto, agora vai vir a presencial. Então, se a pessoa for de outro estado, cidade, ficaria enviável.

Entrevistadora: É quase que um benefício, né? Que algumas empresas dão. Tipo, ah, te dou esse benefício de trabalhar remoto, porque realmente, assim, da nossa área, tecnologia é difícil alguém querer presenciar.

Afonso: Exato.

Entrevistadora: Exato, tipo, ah, essa vaga é só remoto, eu não quero muito complicado.

Afonso: É difícil. Essa é a sua regra.

Entrevistadora: É. Daí, a próxima etapa seria, assim, como é que você vai desenvolver esse algoritmo, né? Você falou dos dados, falou dos requisitos, dos cenários, fez uma reflexão aí sobre como esses dois, essas duas questões se conectam. E aí, eu queria saber um pouco, assim, quais métodos você escolheria para montar esse algoritmo. Se você acha que teria algum desafio técnico, sabe, para conseguir os dados ou para fazer a divulgação e como você faria integração e implementação desse sistema de divulgação.

Essa parte, ela está mais atrelada até, assim, se precisar, ela já até falou, assim, olharia no LinkedIn, já, mas, assim, não precisa ter uma materialização exata de como você faria essa integração. Mas, um pensamento, né, como eu falei, seria assim, você recebeu esse desafio, esse desafio de construir essa solução, você tem que pensar em tudo para ver se vai falar com alguém, se, né, como é que você, se vai precisar de mais gente, enfim. Mas, esse recorte.

Afonso: A respeito, peraí. Deixa eu botar isso aqui também. Vai manter esse slide?

Entrevistadora: Você quer voltar naquele? Eu volto lá.

Afonso: Ah, pode ser, então. Tá. Deixa eu ver aqui.

Entrevistadora: Você pegou essa aqui? Com... Peguei. Tá. Voltei.

Afonso: Então, sobre o desafio técnico. Deixa eu pensar, Larissa. Porque o primeiro desafio, como eu vi ali, é mais, é mais entender como puxar cada tipo de informação no começo, né? Começando um pouquinho antes, assim, tá? É... Tá, então, eu tô na ferramenta, preciso puxar todo tipo de informação ali presente, mas não pode também sair bagunçando tudo, senão isso vai dificultar o meu modelo de entender o que é cada coisa.

Então, primeiro, eu estudaria a plataforma para saber como que está disponibilizar as informações, e separar essas informações com base no seu contexto. Então, tem a parte da formação acadêmica, profissional, mais embaixo ali, tem a parte ali de, ah, fala sobre você mesmo, então, a parte da

introdução. Então, primeiro, eu pensaria nessa questão de como está disponibilizadas as informações ali para quando eu desenvolver algo, algum script, para puxar essa informação, levar isso em consideração.

É... Experiência, primeiro, primeiro estudo, assim, da plataforma em si, da que a gente vai estar integrando ali para depois desenvolver um modelo. Bom, mas, eu entendo também que essa empresa não vai estar puxando um por um as informações das possíveis candidatas. Então, primeiro teria que pensar uma forma de como fazer integração com a plataforma para puxar essas informações, coletar essas informações.

Então, tem certas plataformas que já dão apoio para a equipe de desenvolvimento, fornecendo algumas bibliotecas, algumas APIs, que a gente chama, para consultar ela, botar a palavra-chave, ela retorna tudo bem estruturado, as informações, isso facilita muita coisa. Mas se for plataforma que não tem esse tipo de suporte, isso envolve um outro time, daria para fazer, daria, só que não é a expertise dele, que seria de como puxar essas informações. A gente chama de crawler, isso.

Então, abre crawler. De entrar no site, jogar algumas palavras ali, pegar o retorno da pesquisa e tudo mais. Então, primeiro procuraria entender se as plataformas que a gente vai trabalhar nesse projeto, nessa solução, se tem algum serviço para facilitar essa integração.

Se não tem, já consultaria pessoas com desenvolvimento full-stack, não full-stack, mas back-end. E por se tratar de um web crawler e tudo mais, algo como junior, pleno, já atenderia essa demanda de puxar essas informações. Uma vez puxadas as informações, definiriam onde colocar esses dados, tudo mais, para a equipe que vai desenvolver o modelo, uma equipe de centros de dados, desenvolver alguma recomendação.

Apesar de já existirem alguns, não vejo que é um desafio simples de ser resolvido. Então, teria que ter um cientista de dados pleno, sênior, pra tá atuando na atividade. Mas, eu deixa eu pensar.

Uma vez que eu queria estar com o posto dos dados, ter parte do modelo, que vai me recomendar alguns perfis de quantidade. Aí, é importante esse cientista de dados, também, ter uma comunicação com... Ele precisa ter um entendimento, só que, geralmente, não vai ter esse entendimento assim, mas, teria que estar alinhado junto com... Talvez com um headhunter, algo assim, para entender como ele pesquisaria aquele tipo de informação. Então, o cientista de dados é uma pessoa que for desenvolver a regra de negócio, tem que saber como você faz manualmente.

Você vem aqui, procura assim, pega essas palavras aqui, chaves, tudo mais. Até para replicar esse tipo de pensamento no código, para conseguir

filtrar ali as informações. Até para conseguir ajudar no modelo para entender o que é candidato interessante ou não.

Então, teria que ter esse alinhamento entre essas áreas. Porque esse candidato não vai saber muito bem como correlacionar as informações. Ele vai pensar na cabeça dele, só que talvez, talvez não, alguém já pensou nisso e já encontrou o melhor jeito de conseguir correlacionar as coisas.

Então, teria que passar esse conhecimento adiante, para essa pessoa conseguir aplicar essas regras de negócio. Então, como eu não sou da ciência de dados, não trabalho com isso, fica um pouco difícil de ver a parte técnica de Machine Learning que pode ser utilizada. Eu não consegui responder muito bem essa questão para ti.

Mas eu estou raciocinando até chegar a essa parte que eu vejo que também é importante. E também vai existir se eu tiver algo já implementado. Mas basicamente, participação de dados com alguém no momento back-end ali para puxar os dados daquela ferramenta.

Fazer a integração ali. Caso já não exista. Como o algoritmo foi integrado ao sistema de divulgação das vagas da empresa? Vamos pensar como o algoritmo foi integrado ao sistema de divulgação de vagas da empresa.

Quem divulga as vagas é contratante. Fala assim, tem essas vagas aqui. Mas o headhunter vai procurar as pessoas para que...

Entrevistadora: Essa parte é mais assim, como ser integrado. Eu acho que você já falou bastante como é que você pensaria em realmente puxar os dados ou retornar. É por aí. Mas isso aí que você falou.

Afonso: Isso. Eu imagino que esse modelo vai retornar tipo um score daqueles perfis que foram encontrados com base nas regras de negócio dos dados coletados antes da origem. E isso seria escolhido para essa equipe de headhunter, por exemplo.

Eu levo aquele disco esses perfis que mais se... mais deu match com os requisitos. Esse pessoal entra em contato com eles para marcar uma primeira conversa e entender melhor algumas informações que talvez ficou muito genérico ali na plataforma para ver se realmente se enquadra nesse perfil da contratante.

Entrevistadora: Isso.

E aí por fim são dois momentos aqui que eu acho que são a mesma coisa. A parte de testes e validação. Você já até começou a falar sobre isso nas métricas de sucesso.

Scores e tal. Feedbacks e iterações. Como é que você também poderia ir ajustando esse modelo conforme você vai tendo o resultado retornado.

Você faria algum teste? Teste AB? Não sei. Algum teste que tenha a ver com essa parte técnica mesmo para você testar o funcionamento a varredura desse algoritmo. E aí assim esses dois têm a ver que é assim você pensar na parte de resultados e impactos.

Quando você pensa na métrica de sucesso está pensando já na parte dos resultados impactos como que isso pode impactar na contratação em si. E aí pensando até o final do final da jornada. Você escolheu os dados que você vai trabalhar fez o modelo ele rodou e aí retornaram resultados e esses resultados no final dessa jornada eles vão converter em contratações ou não.

Depende de pessoas específicas quem seriam essas pessoas que seriam contratadas. Se tem alguma coisa que você pode inclusive e balizando ele no meio quando você está fazendo desenvolvimento modelagem.

Afonso: Então vamos lá sobre a métrica de sucesso.

O quanto está sendo assertivo o modelo. A contratante com essa equipe que toma a frente para conversar com os candidatos e marcar o dia da entrevista ela fica consciente de a pessoa compareceu teve entrevista o que a pessoa achou a contratante achou né. Então existe algumas etapas do processo seletivo que se essa equipe que fica à frente o Headhunter ele tem ciência.

Então poderia ser muito bem existir um sistema ali dentro não da contratante da empresa que está procurando dele reportar como que foi não somente por texto escrevendo mas se foi um candidato que realmente deu match à empresa ou não. Porque bem ou mal ele para entrar em contato com esses candidatos que se baseou no resultado do modelo. Então passou uma listinha por exemplo vou falar com uns chutar números top 20 com o score mais alto que o modelo retornou vou falar com esses 20.

Então com base no que a contratante foi falar e com base também na primeira conversa que ele teve com o candidato então tem aquela pré-conversa ali ele sabe se realmente aquela pessoa faz sentido ou não. Então isso é uma certa forma de validar o modelo. Então você você armazena essa informação se aquele cara fez sentido ou não anota isso no sistema e tudo mais e isso vai servir como como uma validação do modelo para saber se precisa ser retreinado e tudo mais.

Então a métrica seria a partir disso que eu vejo para analisar o desempenho do algoritmo. E uma vez atendendo baixo tá então é o que você iria pensar por que o modelo está indicando essas pessoas por que essas pessoas estão com o score alto. Aí tem técnicas de explicabilidade do Master Learning que te ajuda a você entender como chegou naquela resposta.

Ou talvez seria um estudo relacionado a isso para entender para explicar aquele resultado que o modelo entregou para não ser uma caixinha preta.

Esse era um problema muito antigo no passado que agora está começando a melhorar com essas novas técnicas de explicabilidade. E aplica e a partir dessas técnicas você tem buscar entender por que chegou naquele resultado e fazer as devidas alterações.

E é melhor o jeito que eu vejo e também na hora de treinar o modelo também uma coisa importante, é assim além de além de você poder coletar manualmente alguns candidatos que você entende para aquele perfil e ver como que o modelo reage com essa com essa série de candidatos. Que provavelmente não entrar no treino senão acaba manipulando o resultado não fica viciado. Você pode utilizar isso como validação do mesmo modelo.

Então você sabe que o conjunto de pessoas ali, candidatos atende hoje vai ser aquele modelo vamos ver como ele vai retornar a informação. Então seria uma técnica também a ser feita. Como avaliar os feedbacks internos e externos que podem levar a manipulação dos algoritmos? Bom, aí uma vez que a empresa a contratante contratou uma empresa para fazer essa parte inicial de filtrar as pessoas para participar do processo seletivo seu e você vê que você encontra vários candidatos que não levam a lugar nenhum em lugar algum com certeza essa empresa vai tomar uma atitude quanto a isso ou troca de empresa para procurar candidatos melhores ou continua com ela e reposta a informação para eles.

Então com base nessas informações a empresa também sabe se está conseguindo atingir os candidatos certos para cada contratante. Então isso também é uma boa métrica uma boa forma de avaliar o rendimento do modelo. Se uma vez você está conseguindo manter aqueles contratos com a contratante por questões você está conseguindo entregar bons candidatos para ela é uma forma também que faz você opa então o modelo que a gente está utilizando para conseguir filtrar os candidatos é algo que está sendo diferencial no mercado tanto que as empresas estão mantendo um contrato comigo novas empresas estão chegando e estão gostando da possibilidade do modelo formulário que a empresa pode preencher a contratante isso tudo ajuda a avaliar o modelo de recomendação.

Mais o que? Quais tipos de testes seriam realizados para validar o algoritmo? Bom, eu já disse mais ou menos alguns testes algumas formas de validar você pode ter alguns cases certos de candidatos que tem aqueles requisitos e ver se para esses testes estão batendo e outra forma indireta é essa questão de começando os feedbacks seja da contratante seja dos próprios headhunters ao conversar pela primeira vez com os candidatos para saber se está fazendo sentido ou não. Pode avançar o slide um pouquinho? Deixa eu ver.

Entrevistadora: Acho que você também comentou, né?

Afonso: É, acabei comentando

Entrevistadora: É um pouco junto esses dois, sim.

E deixa eu te perguntar em relação às pessoas que de fato vão ser buscadas vão ser vão passar por essa peneira você considera que você teria algum direcionamento para fazer para eliminar algum viés alguma coisa assim?

Afonso: Pode me fornecer um exemplo? Ou explicar de outra forma a pergunta?

Entrevistadora: Sim. Por exemplo você vai fazer ali lá no começo, né? Pegou os requisitos entendeu melhor vai começar a fazer as regras de negócio fazer lá o filtro dos dados enfim você pode selecionar assim do jeito que você vai montar esse filtro ele pode acabar retornando pessoas muito parecidas assim por exemplo porque é uma bolha, né? Assim uma bolha das pessoas estão ali fazendo desenvolvendo trabalhando com determinadas ferramentas específicas e tal na hora que você vai fazer um algoritmo, né? Vai pensar assim como fazer montar esse modelo você acha que você precisa ou deveria né? ou ia querer fazer algum filtro ou fazer algum direcionamento para eliminar algum possível viés que venha a ser a criar que venha a ser criado?

Afonso: Deixa eu pensar um pouquinho Larissa tá muito difícil de pensar

Entrevistadora: Você entendeu? Se não entendeu eu repito tá de outra forma

Afonso: Acho que a gente ia sim Agora eu estou pensando como que eu faria Um momento Então a gente tem a lista de digamos de esportes e as quantidades que mais dão match com a empresa Tá O que acontece nessa lista o próprio Headhunter ele pode ter algumas

(FIM DA TRANSCRIÇÃO)

B.2

Bruno

Entrevistadora: Deixa eu só mudar o... Ah, já veio aqui, tá em português, tá ótimo. A minha ideia, o que eu pretendo, né, é que a gente tenha uma conversa, eu chamo de entrevista semi-estruturada, porque tem alguns tópicos que eu vou trazer, eu vou até compartilhar para ficar mais fácil você vendo, né, e assimilando que ponto que a gente está, tá, mas a ideia é que, assim, é uma simulação, né, eu quero analisar mesmo o seu modo de construção do pensamento, assim, quais decisões você tomaria diante de uma situação hipotética que eu vou te passar, tá, o que realmente o meu interesse aqui é entender quais são as suas linhas de pensamento, como você estrutura mesmo,

estruturaria essa solução que eu vou te passar como uma solução hipotética, tá certo?

Bruno: Tá certo.

Entrevistadora: Vou compartilhar aqui, você tá vendo minha tela?

Bruno: Estou.

Entrevistadora: Tá, então, a princípio, eu gostaria de entender, assim, até para depois conseguir contextualizar melhor, ó, um pouco do seu papel na empresa que você tá e a sua experiência com algoritmos, né, para entender, assim, o que você já passou, um pouco da sua experiência profissional e no todo, e se você puder fazer um recorte também nessa questão de se em algum momento você trabalhou com algoritmos, seria legal.

Entrevistadora: Então, Larissa, nesse exato momento eu estou como gerente aqui na Deloitte. Eu já entrei na Deloitte como consultor sênior e, na verdade, mais trabalhando como cientista de dados, especificamente na parte da empresa computacional, já trabalhei com alguns sistemas de machine learning, da parte mais modelos preditivos, lá no início, depois passei para já modelos de visão computacional, fiz várias coisas com algoritmos de visão computacional aqui na Deloitte, vários projetos, e, bom, e depois que fui promovido para gerente, tenho feito algumas coisas, mas já aí como não foco na minha função, não são tantas assim, não.

Lá no passado, já venho trabalhando com algoritmos de visão computacional, processamento de imagens, desde aproximadamente 2007, já fui entrando mais na parte da inteligência artificial, machine learning, a partir de 2008, com minha tese de doutoral, e a partir disso, bom, fiquei esse tempo todo trabalhando projetos, desenvolvendo algoritmos de visão computacional para clientes, enquanto estava ainda na universidade, que é no caso na PUC Rio, onde eu fiz o meu doutorado e depois fiquei, faço um pós-doutorado lá, até que entrei em 2019 numa empresa, uma startup, que também prestava serviço para outras empresas de inteligência artificial, visão computacional, etc, parecido com a Deloitte, só que uma empresa menor, e ali trabalhei vários projetos com várias empresas, também desenvolvendo algoritmos de visão computacional, resolvendo problemas nos projetos, enfim, soluções e tudo mais. E, bom, já da Deloitte, já falei.

Entrevistadora: Tá legal, muito bom.

Eu vou te passar aqui o contexto, né, desse objetivo e te explicar, se você tiver alguma dúvida, sempre pode parar, perguntar, tá bom? Vamos dizer que uma empresa de tecnologia de porte médio vai contratar uma outra empresa para fazer uma divulgação de uma vaga que está aberta para anunciar a seguinte oportunidade, né? Uma vaga de tecnologia para

profissional pleno, que seja para um desenvolvedor ou desenvolvedora full stack, não passou muitos detalhes, que tenha formação em engenharia de software, tecnologia da informação, sistemas, afins, ou seja, áreas de rearia e ciências, e uma experiência de cinco anos de mercado, né, mais ou menos, e um regime remoto. E aí você é o profissional que está na empresa de divulgação de vagas que vai ser contratado. Então, a empresa de tecnologia vai te contratar para desenvolver um algoritmo, um algoritmo de recomendação ou algum outro tipo de algoritmo, para você formatar esse anúncio com essas características específicas para buscar chegar até pessoas que tenham o perfil, que compreendem essa vaga, que dão o match, né, que dão ali o fit com essa vaga, com esses requisitos que foram anunciados, tá? Tá.

E aí, assim, primeiro passo, como você trabalharia na definição desses requisitos? Tem alguns requisitos ali que foram passados, mas como você consegue, de repente, elaborá-los mais? E quais seriam os dados utilizados? Como você buscaria esses dados?

Bruno: Dados, você está se referindo, pesquisar sobre pessoas que têm fit com aqueles requisitos, é isso?

Entrevistadora: Isso, sim. Mas como você transformaria em dados, né, assim, para você usar a característica de cada uma dessas pessoas? E aí, assim, a gente pode fazer um exercício aqui de pensar, imagina, desenvolver um anúncio no Facebook, aqueles anúncios que ficam aparecendo. E aí, aquele anúncio, ele aparece de acordo com a sua, seu estilo, né, suas preferências, algum dado lá que você já gerou, que ele vai conseguir analisar seu perfil.

Então, a ideia era, assim, como você consegue tirar esses dados para direcionar esse anúncio, né, que vai ser anunciado, né, que vai ser divulgado.

Bruno: Tá. Primeira coisa, eu procuraria por hashtag ou palavra-chave sobre formação.

Essa pessoa teria que ter algum tipo de formação com TI. Outra coisa que eu procuraria é tempo de também procurar palavras-chave, onde essa pessoa tenha relatado, tenha colocado ali. Porque, normalmente, quando você cadastra nessas plataformas, essas redes sociais, seja LinkedIn, por exemplo, Facebook, ou for, você preenche ali num histórico de onde você trabalhou, quanto tempo e tudo mais.

Então, eu tentaria procurar quanto tempo essa pessoa ficou alocado naquelas funções relacionadas com TI. Se tem, também, se a formação, quanto tempo está formado, também é importante. Também procuraria nos posts históricos, coisa que ele tenha escrito, dependendo de qual for a rede social, relacionado com o assunto, para saber, mais ou menos, o interesse dele, quão ativo essa pessoa é nesses tópicos, nesses assuntos.

Eu não lembro muito bem. Larissa, você tem como ir, por favor? Claro. Por enquanto, eu estou... Acho que seriam as coisas que eu procuraria.

Eu estava com dúvidas se tinha esquecido alguma coisa, mas acho que seria os tópicos que eu tentaria procurar para conseguir encontrar as informações necessárias para cada possível candidato para essa vaga. Se eu fizesse um algoritmo de recomendação, alguma coisa do tipo, eu tentaria focar nessas coisas e alimentar ele com esses dados.

Entrevistadora: Tá certo.

E aí, você, nesse contexto, você... Qual método você escolheria? Ou uma técnica de machine learning, ou um método diferente para o algoritmo específico? Se você acha que teria algum desafio técnico que você consegue mensurar, consegue imaginar que pode enfrentar? E aí, essa parte de integração e implementação, você pode chegar, né? Se você tiver alguma ideia, você pode dizer, mas não é necessário, tá? Assim, é mais, vamos dizer, se alguma pessoa tiver má experiência e já tenha feito isso com alguma plataforma, alguma rede social, mas assim, não é necessário, tá? Porque a gente não está analisando exatamente a efetividade da técnica, mas sim a construção do pensamento.

Bruno: Tá, então, eu... Primeiramente, eu teria que ter um algoritmo de busca exaustiva, logicamente, para procurar dados. Vamos supor que já o algoritmo está treinado, tá? Qual algoritmo eu treinaria? Não sei, pode ser um algoritmo de classificação.

Depois que eu tenha o meu algoritmo treinado, eu iria procurar com um algoritmo de busca exaustiva, de currículo, iria procurar pessoas com aquelas características, baseado em uma classificação que eu vou fazer, baseado nos parâmetros que eu defini lá, as features que eu defini para esse meu classificador. Seja, não sei, tempo de formação, formação, tipo de formação, tempo de empregado, não sei. Teria alguma dessas features, tá? E aí, trataria de... Logicamente, se for um algoritmo supervisionado, eu já vou treinar nele com um conjunto de treinamento, validação, etc.

De forma tal de que eu já saiba quais são as médias do meu algoritmo e tudo mais, eu já sei qual vai ser a confiabilidade do resultado dele previamente, tá? Então, eu já sei que o que ele me entregar ali, ele vai ter um nível de confiança X, sobre aquela superfície eu vou poder confiar com aquelas atividades. Não sei se é o algoritmo mais recomendado para isso ou não, mas talvez seja o que eu falaria. Desafios técnicos.

Eu acho que o mais desafiador aí não é a implementação do próprio algoritmo, eu acho que seria a extração de todas essas features que eu vou usar para alimentar ele. Você vai ter que ter um sistema, como eu já falei, de buscas auditivas para o currículo extrair, definir quais são aquelas palavras-chave que

vão ser usadas, que seriam as features, no caso. Você tem que ser capaz de encontrar nelas em qualquer um dos currículos, ou boa parte deles, para você ter features suficientes para em todos os casos.

Talvez você saiba que a pessoa nem preenche corretamente os dados na hora que se cadastra nas redes sociais, nem todas as informações estão lá. Então, escolher aquelas que tenham a menor quantidade de vacâncias possíveis, ou seja, aquela que às vezes se espalha mais, que a pessoa não preenche, procurar aquelas informações específicas, achar aquelas informações específicas, é uma das coisas que eu acho mais desafiador aí. Depois que você tem isso, depois que você tem os dados, eu estou treinando o algoritmo.

Aí já não vejo muito desafio. Então, interação. Se você já tem um sistema, um modelo já treinado, é simplesmente fazer deploy dentro daquele algoritmo que existe nessa plataforma.

Vamos supor, o LinkedIn. O LinkedIn já tem modelos próprios que estão rodando dentro da plataforma e para outros objetivos, não é precisamente isso, mas vamos supor que o LinkedIn me pediu para criar um modelo para pesquisar sobre, para ajudar uma pessoa de RH, na hora, um headhunter, que esteja, vamos criar aqui um serviço para headhunter, para facilitar a vida deles, deve encontrar pessoas com certa característica. Então, eu, logicamente, teria que fazer um serviço dentro do LinkedIn, dessa oportunidade da pessoa colocar os dados que ele quer da vaga, ali, e esse algoritmo vai procurar match.

Graças ao seu treinamento, estou partindo do ponto de que o modelo foi treinado, foi um treinamento supervisionado. Estou partindo do ponto de vista que seja um não supervisionado, mas poderia ser não supervisionado também. Vamos supor que o treinamento seja supervisionado, então o modelo já foi treinado, pensando todas as variações possíveis.

Então, a pessoa vai preencher ali seus dados, aquele modelo vai ali procurar, fazer aquela busca do seu objetivo e já com um algoritmo próprio do LinkedIn. E aí ele vai separar o que ele acha que dá match com aquele característica, aquele requisito que a pessoa colocou e aí vai devolvendo um resultado para essa pessoa. Mas tudo isso também poderia ser feito com um modelo não supervisionado.

Aí seria mais simples o ponto de vista de que você não precisa fazer treinamento e tudo mais, mas nem sempre os resultados são tão bons quando com um modelo não supervisionado. Aliás, não supervisionado, isso aí.

Entrevistadora: Tá bom.

E aí assim, em relação a teste, e já pensando nas métricas de sucesso, lembrando que assim, você foi uma empresa que foi contratada, você precisa mostrar para o seu cliente que esse algoritmo está funcionando, está tendo bons

retornos e pensar nessas métricas para medir a assertividade. E os feedbacks que você pode coletar, tanto humanos como não humanos, para você fazer iterações.

Bruno: Teste para validar no algoritmo, acho que é fazer simulações.

Você teria que fazer simulações ali com... pesquisando sobre, preenchendo ali aqueles requisitos que você quer sobre a pessoa e aí ver no que aquele algoritmo está te devolvendo. Como resultado, ver se realmente aquilo que ele está te devolvendo te satisfaz. Eu gostaria, sinceramente.

Agora, métricas. Eu usaria as métricas que normalmente uso em qualquer modelo. Tentaria ver qual é a precisão do modelo.

O recorde. F1 score, por exemplo. Eu não usaria a acurácia, nunca gostei da acurácia.

Meio pesado. Então, por exemplo, precisão. Eu tentaria... Primeiro, eu tenho que ver quais são os fases positivas, as fases negativas e o verdadeiro positivo que eu tenho.

Então, vamos supor que meu modelo me devolveu dez resultados na pesquisa. Nesses dez resultados da pesquisa, eu tenho cinco pessoas que se encaixam perfeitamente. Eu tenho, sei lá, três pessoas que não têm nada a ver com aquilo que eu estava procurando.

Aliás, nesse caso, enfim. São dez pessoas, eu tenho sete que são verdadeiros positivos, porque realmente tem tudo a ver. Eu tenho três pessoas que não têm nada a ver com aquilo que eu coloquei.

Então, nesses casos aí, eu diria que eles são falsos positivos. E vamos supor que no meu banco de dados, eu sabia que tinha exatamente dez pessoas que tinham MAT. Então, tem três pessoas que são falsos negativos que não foram incluídas aí.

Então, ele incluiu três falsos positivos, porém deixou de incluir três. Então, esses outros três que ficaram de fora são falsos negativos. Então, por isso aí, eu já saberia qual é a minha precisão, que seria o verdadeiro positivo dividido pela soma do verdadeiro positivo mais falsos positivos.

Eu saberia qual é o meu recorde, que seria o verdadeiro positivo dividido pela soma do verdadeiro positivo com falsos negativos. Aí já saberia qual é o precision recorde do meu modelo, e eu ia fazer uma escolha, dois vezes o produto do precision recorde dividido pela soma deles. Então, essas seriam as métricas que eu usaria, só que normalmente eu uso o modelo, porque com isso você sabe se teu modelo está... Ou seja, indo em que caminho, se você está com muito falso positivo ou se você está com muito falso negativo.

Ou seja, se você está achando que muitas pessoas se encaixam quando não se encaixam, ou deixando passar pessoas que não estão sendo pegadas,

porque elas estão passando pela peneira e você não está conseguindo enxergar nelas, entende? E eu faria uma escolha com uma acurácia, que finalmente é ponderada ali no precision recorde, que faz um papel de acurácia, porém ele é menos embeçado. Então, eu usaria essas métricas. Feedback e interações.

Eu colocaria nesse sistema, logicamente, para esse retorno, por exemplo, ele dá um feedback sobre como ele achou a ferramenta, e como a ferramenta foi útil para ele. Porque, por exemplo, para eu poder gerar essas métricas, eu teria que ter conhecido, ou essa pessoa teria que ter me falado, das 10 pessoas que me sugeriram, 7 tiveram match perfeito e 3 não tinham nada a ver, por exemplo. Só que isso é difícil, ter esse feedback.

Então, isso aí você, essas métricas que eu estou falando assim, dificilmente um algoritmo desse você vai ter. Que você pode usar no sistema de feedback com a pessoa ali, e a pessoa preenche. Eu tive, sei, tantas pessoas recomendadas, dessas tantas, apenas uma 30

E isso, eu poderia tentar refinar um pouco mais esse feedback, para saber em que sentido eu teria que, quais são os parafusos que eu teria que apertar no meu modelo para melhorar a assertividade. Ah, eu estou focando muito naquelas, aquela palavra chama, aquelas features específicas, e eu estou deixando passar outras coisas que são mais relevantes. Enfim, e aí eu poderia refinar meu modelo e tentar melhorar o resultado.

Entrevistadora: Nessa parte, você considera que seria, como a gente está falando assim, seleção de pessoas, né? Você considera que seria importante ter uma troca com, sei lá, o cliente, alguém especializado em RH, ou alguém que tenha essa visão mais da sociedade em si, né?

Bruno: Eu considero bem importante essa interação, porque só dessa pessoa você vai ter um resultado prático, ou uma ideia prática de como aquele algoritmo está se comportando. Porque uma coisa que você testa na hora que você faz a modelagem ali, está criando um modelo, está fazendo o teu algoritmo e tudo mais, embora você tenha as etapas ali de validação, testes e tudo mais, isso não é um ambiente real, não é uma produção. Quando você colocar na produção, você vai ter uma viabilidade muito maior de problemas, de situações, que se saem quase sempre daquilo que você tinha, aquele ambiente mais controlado ali.

Então, o único jeito que você tem de receber feedback desse tipo é através de um sistema de interação com o cliente. No final, porque quando você está ali no teu sistema de desenvolvimento, você no prazo do treinamento superficial, vou te repetir, você tem um trabalho prévio onde você teve que fazer toda a anotação, separação entre a classe e tudo mais, então ali você sabe o que está certo e o que está errado. E você vai treinar teu modelo baseado naquilo ali.

Depois que você sai desse mundo, esse mundo tem uma variabilidade relativamente restrita, porque você está ali com uma população de dados um pouco restrita, logicamente. Quando você sai para o mundo que é aí, a quantidade de dados é muito maior, a variabilidade dele é bem maior também, é que você começa a ter os resultados reais, mas você já não tem aquela classificação, aquela separação, aquela anotação que você tinha ali no teu ambiente controlado onde você fez o treinamento e implementação do modelo. Ali, então, você precisa de alguém que te fale, aquele caso está falhando, esse caso está acertado.

Então, esse sistema de feedback é essencial, eu considero. Opa, eu perdi uma parte que deu uma travada, eu não sei se você explicou melhor, você falou que você não gosta muito de acurácia, não gosta de medir o sucesso pela acurácia também. Eu não gosto muito da acurácia pelo fato de que ela é um pouco enviesada, então ela pode dar um resultado que não é muito condicente com a realidade, eu confio mais no resultado do F1 Sport.

Eu tomo o F1 Sport como se fosse uma acurácia, e ela é mais realista que a acurácia em si, entendeu? Então, eu normalmente não uso acurácia. Se você tiver dados, por exemplo, você tem um conjunto de dados onde você, uma porção desse teu conjunto de dados tem uma variabilidade muito grande e eles estão puxando os mapas para um lado que para o outro, tua acurácia pode ser muito boa ou muito ruim, quando na verdade não é o caso. Ela pode ser muito influenciada por conta disso, é como, por exemplo, eu pego um elevador agora com Elon Musk, eu faço a média disso e eu posso falar que no elevador tem duas pessoas mais ricas do mundo fazendo a média do dinheiro que eu tenho e que ele tem.

Isso é meio insignificante, a média não vale muito a pena porque o valor dele é muito grande. Então, a acurácia peca por conta disso. No final, ela usa o verdadeiro positivo dividido pela soma do resto, o verdadeiro positivo e o verdadeiro negativo.

Então, isso no final, o resultado não é muito confiável. Dependendo de como os teus dados são, ela pode te dar a sensação de que você está no caminho certo, você atingiu um resultado muito bom, contrariamente ao resultado muito bom, ou vice-versa. Então, normalmente eu uso o F1Score no lugar da acurácia.

Entrevistadora: Legal. Entendi. E aí, assim, eu acho que essa parte final, acho que você já até comentou um pouco, trouxe nessa parte da efetividade em si, né, para avaliar como o algoritmo está realmente dando fit nos candidatos que são mais recomendados para a vaga e o impacto na contratação. Acho que você já falou um pouco disso, mas não sei se quiser complementar.

Bruno: Eu acho que para você avaliar efetividade, efetividade em que

sentido? Se a pessoa foi contratada ou não foi contratada?

Entrevistadora: Não é nem exatamente isso, pode ser uma métrica da contratação ou não, mas se realmente o que você modelou está encontrando pessoas que tenham a ver, né? Assim, acho que você já falou bastante até de algumas técnicas que você colocaria para medir os falsos positivos e os negativos, né?

Bruno: Acho que já por aí já você teria uma ideia de quão efetivo é teu modelo. Agora, isso também, como eu falei, é no ambiente de desenvolvimento.

Quando você coloca isso na produção, você não vai ter os falsos negativos, os falsos positivos, a não ser que alguém te fale. No meu caso...

Entrevistadora: Só um minutinho, só um minutinho. Oi, perdão, eu tinha pedido água e o cara chegou no meio, não. Perdão, pode continuar.

Bruno: Não, tudo bem, não tem problema, não. Então, o que eu estava falando é...

Entrevistadora: Cortei sua linha de raciocínio.

Bruno: Não, tudo bem. Então, o que eu estava falando é que uma coisa são os falsos positivos, falsos negativos, verdadeiros positivos, em um ambiente de desenvolvimento onde você conhece os dados, você já separou os teus dados e aí você pode fazer esse cálculo. Outra coisa é quando você coloca aqui na produção.

Quando você coloca aqui na produção, normalmente, você não vai lá e classificar as coisas. Você não tem tempo para isso e normalmente são muitos dados. Então, se você tiver um sistema ali onde o cliente der um feedback, o que você poderia usar para você ter uma ideia de quão bem a sua ferramenta está se comportando.

Então, uma coisa é o que você tem na hora do desenvolvimento, que é o que eu colocaria como métricas, e outra coisa é o que você tem quando tem produção. Então, você tem aquele sistema de interação com o cliente, o feedback, que eu acho que é bem interessante e bem importante. Normalmente, por exemplo, as redes sociais, o YouTube, eu uso as joinhas.

Se você dá joinha que gostou, se você dá joinha para baixo que não gostou, você tem as estrelas das pessoas na estrela. Aqui tem nota máxima, cinco estrelas, ou tem menos estrelas. Enfim, tem que ter alguma forma de o cliente, a pessoa, te indicar se gostou do produto, se o produto foi útil, quão eficiente foi o produto.

Ou você precisa de um feedback de algum tipo, dessa forma desse cliente. E aí, a efetividade desse modelo vai ser baseada nesse feedback. Então, se eu receber, vamos supor que você tem estrelinhas, se eu receber quatro estrelas em

média, tem quatro estrelas em média, então, você sabe que a sua efetividade é muito boa.

Entende? Então, você tem que ser baseado em algum tipo. A pessoa tem que dizer, ah, isso aqui funcionou para mim. Isso não está funcionando, isso está ruim.

Entende? E aí, você pode ter algum inclusivo, alguns parâmetros a mais para ele te ajudar. Aí, botamos a coisa do lado anterior. A pessoa está ajudando você a entender onde você está errando, onde o seu modelo está se fechando ou se perdendo, deixando de entregar o resultado que ele deveria entregar.

Para você, então, retreinar e você fazer aqueles ajustes, aqueles parâmetros, ou talvez no teu dataset, etc. Para você conseguir melhorar o resultado. Entende?

Entrevistadora: Entendi.

Bruno: Então, pode falar. Não, isso aí era a parte da efetividade e tinha outro ponto que seria o impacto na contratação. E assim, essa aí também é outra coisa que acho que seu cliente pode ter falado.

Entende? Que acho que pelo feedback também você entenderia. Se realmente esse sistema aqui me ajudou a contratar pessoas que eu queria, logicamente que eu vou te dar uma nota boa. Entendeu? Se não, eu vou... Então, isso tem que ser com uma interação com o cliente.

Você tem que receber o feedback para você saber disso. Quando você sai da parte do desenvolvimento, você não tem mais essas informações. Você precisa que alguém tenha falado.

Entrevistadora: Eu acredito, então, que uma equipe multidisciplinar, com as visões diferentes, ajudaria nisso. Tivesse atuando não só na construção, como também na avaliação ali do desenvolvimento.

Bruno: Certamente.

Você tem que, além de você ter um especialista em TI para fazer toda a parte do desenvolvimento, do modelo, as pessoas que vão saber quais são os algoritmos, os melhores algoritmos para você usar ali, o classificador estar com tais parâmetros. Enfim, tem que ser uma pessoa especialista em aquilo ali, beleza. Mas, além disso, essa pessoa precisaria ter um conhecimento já de... Headhunter, por exemplo.

Provavelmente, essa pessoa não tem. Porque quando você é especialista em TI, você normalmente está focado naqueles algoritmos matemáticos e tudo mais. Mas você não tem visão de uma pessoa de RH, por exemplo.

Ou de um Headhunter, por exemplo. Essas pessoas já são treinadas, elas sabem o que procuram, elas veem o currículo e já, de cara, já sabem quando uma pessoa está dando o máximo ou não com aquilo que ela está procurando.

Então, elas têm já aquela experiência, aquela expertise própria que dificilmente alguém que está lá na parte do desenvolvimento, o modelo, tem.

Então, essa pessoa, logicamente, precisaria estar trabalhando junto também. Por exemplo, a parte ali da extração das features, no início que eu te falei. Ah, quais são as palavras-chave que eu vou procurar? Quais são as coisas que eu vou procurar aqui para treinar meu modelo? Essa pessoa é que tem que dizer, entende? Ah, quando eu vou fazendo isso, eu olho nesses tópicos aqui.

Vejo se a pessoa se formou, qual foi a universidade, que tipo de universidade, quanto tempo ficou para terminar na graduação, entende? Ah, se ele fez a graduação em tempo, não fez a graduação em tempo, quanto tempo se formou. Depois que se formou, ele foi automaticamente para trabalhar ou ele demorou para entrar no mercado de trabalho. Quanto tempo está, entende? Tem algumas macetas, algumas coisas que provavelmente essa pessoa, estou aqui falando com esse buraco porque não tem essa experiência, mas deve-se ter algumas macetas que essa pessoa, treinado já, conhece.

Então, ele vai direcionar para essa pessoa, sabe, especificamente aqui. Porque, do contrário, você vai ter que classificar o que vai classificar, só que não vai classificar aquilo que o cliente quer, entende?

Entrevistadora: Entendi. Sim, e você acha que assim, nessa parte, quando a gente fala dessa interação multidisciplinar, você acha mais importante a temática, o conteúdo da pessoa? Você considera também a questão de diversidade, de colocar pessoas diferentes, de perfis diferentes, ou gêneros diferentes? Você considera isso geralmente?

Bruno: Sim, é importante, até porque você está desenvolvendo uma ferramenta que pode ser usada por diferentes clientes.

Você pode ter uma ferramenta para uma empresa que tem uma cultura X, onde eles são, não sei, gostam de ser inclusivo, em todo sentido, em gênero, com, se tratar, mas você pode estar desenvolvendo isso aí para outra empresa em que eles estão mais focados em alguém que realmente seja competente e não vão estar olhando muito para essa parte. Então, você tem que abranger a maior quantidade possível de alternativas, entende? Então, quanto mais você tiver de pessoas que olhem para todos esses pontos, você vai ter um resultado lá que, na média, é melhor. Entendeu? Você vai mais satisfazer, de forma geral, seus clientes.

Então, acredito que diversidade, ali na hora de você ter critérios diferentes na hora de você fazer as escolhas, é importante para que o seu modelo fique o mais generalista possível. A ideia é que, quando você cria um modelo, ele generalize bem. Ele não fica muito focado em, ah, eu só vou focar na pessoa que é muito boa em aquilo que o cliente está procurando.

Entendeu? Também quero abrir mais o lake para pegar também outras coisas que outros clientes poderiam gostar também.

Entrevistadora: Entendi. É isso, Bruno.

Muito obrigada. Foi muito legal. Você trouxe muitas questões técnicas importantes e que conseguiram até me clarear melhor algumas possibilidades de estratégia para essa solução.

Mas foi muito importante. Obrigada, viu?

Bruno: Nada, imagino. Eu fico feliz em ter ajudado.

Entrevistadora: Muito bom. E desculpa os contratempos, tá?

Bruno: Nada, imagino. Tchau, Larissa.

Entrevistadora: Tchau, tchau.

Bruno: Sucesso aí. Tchau, tchau.

Entrevistadora: Obrigada.

B.3

Caio

Caio

Entrevistadora: Que já ajuda bastante, mas deixa eu ver se tá em português. Sim, tá em português. Então tá jóia.

Eu vou aqui compartilhar a tela, caso você queira compartilhar a sua tela, mostrar alguma coisa, material, alguma ferramenta, fica à vontade, tá? Só pegar aqui o acesso. A princípio então, como eu te falei, né? A gente quer investigar aqui os métodos que você escolhe, seus caminhos, seu mapa mental para fazer um desenvolvimento de um algoritmo. E aí, a princípio, eu gostaria que você fizesse uma contextualização falando um pouco da sua experiência profissional com algoritmos e da sua experiência atual.

Caio: Eu vou passar um pouco pela minha carreira e minha formação para chegar aqui. Minha formação é em estatística e durante o curso eu fiz duas cadeiras principais de algoritmo. Tenho outras que foram introdutórias, postalzinho, mas eu tinha R e tinha SAS, né? As duas linguagens.

Fui passando por esses softwares, utilizando esses softwares na faculdade. Quando eu chego para estágio, eu uso principalmente R. No meu primeiro trabalho, Dupre, eu desenvolvia, basicamente, o relatório, o relatório gerencial. Então, eu precisava montar o algoritmo para filtrar alguma base, para ajustar a base.

Basicamente, era esse o processo.

Entrevistadora: Tá legal.

Caio: Eu saio da Dupre e vou para a Sul América. Aí, eu já começo a carreira no mercado de seguros, que eu conto até hoje. Sul América, eu

começo a fazer modelagem. Já é um pouquinho mais diferente da questão de desenvolvimento. Continuo fazendo algoritmo para base de dados e ajuste de base de dados e limpeza. ETL no contexto em geral. E eu começo a trabalhar com modelagem em si.

Aí, a modelagem, eu já não desenvolvo o algoritmo, já uso um algoritmo já desenvolvido. No mercado de seguros, tem um software que é muito utilizado, que é mais fixo, que é o Angular, que facilita os ajustes das variáveis. Quando sai essa variável resposta, a estimação da variável resposta pelos modelos, a gente tem um Excel com os fatores de cada uma das variáveis.

As variáveis para entrar na empresa, elas sempre entram categóricas. Então, se eu tenho uma variável numérica, a gente categoriza ela e acha para entrar com essa variável e ter lá um fator, um parâmetro associado a cada nível dessa variável. E eu fiz um algoritmo para decodificar isso, para poder fazer com que a nossa estrutura de cotação, o contador da empresa, lesse esse modelo.

Então, tinha um algoritmo pegado. Era um IFELS, mas eu tinha que usar um outro tipo de linguagem. Era uma linguagem RTDM, que já é um produto SAS, mais específico assim, mas a estrutura é basicamente de IFELS.

E desenvolvimento de algoritmo para relatórios e acompanhamentos, que são mais gerais. Eu vou pular a Allianz e vou para a Slippage, porque a Allianz foi uma compra, foi o MA da Sul América. Então, o processo era muito parecido, a mesma coisa. Quando eu chego na Liberty, eu já tenho uma outra estrutura. Eu fui para a equipe global da Liberty, então ela não estava fazendo parte do Brasil. E a equipe global tinha muito uma experiência de pesquisa.

Então, ela mais pesquisava do que necessariamente implementava ou fazia esses processos que ela fazia nas outras seguradoras. Então, eu tive uma experiência muito forte de mesclar metodologias e algoritmos e soluções. Temos um modelo de Machine Learning, como que a gente consegue ajustar a variável, ajustar o modelo para fazer com que isso faça parte do mercado segurador.

Eu tive um processo de adaptação de algoritmo, pega determinado conceito, beleza, mas determinada restrição não pode acontecer. Então, a gente junta com outro conceito. Meio que é criando metodologias novas, teorias e mesclando alguns conceitos, que fazem muito sentido para o mercado segurador.

Então, eu acho que são essas as coisas que eu mais gosto. E a partir dali, a partir dali antes, na verdade, em Python, né? Aí já foi linguagem Python que tocou os hoje também. Nessa transição para HD. Então, foi no outro MA.

Entrevistadora: Só uma dúvida que eu fiquei, essa parte das seguradoras,

né, que você fazia, você participou, você tinha, falou que já trabalha com o algoritmo mais pronto, só vai regulando as features, né? Você, essa parte, você faz as decisões sozinho, tem uma equipe, tem, vem de algum lugar da área de risco, alguma coisa?

Caio: É tudo nosso, é tudo dentro da equipe. A gente tem uma influência forte, né, por analista, por quanto o histórico do analista no mercado conta, essas coisas para ajustar. Só que a gente tem um processo que a gente chama de peer review e managing review. Então, a gente desenvolve o modelo, a gente tem as estruturas, né? Tem estrutura para desenvolver, para poder tentar padronizar exatamente, para tentar tirar o filhagem do analista, né? O que o analista pensa, o que ele faz. A gente faz estruturas de, como se fossem etapas, assim.

Primeiro a gente faz isso, depois a gente faz isso. Então, no primeiro processo, a gente faz um feature selection, mas não é bem um feature selection de machine learning, porque o JLM é uma outra estrutura. Então, há mais um, a gente chama de, de, a Pariser aqui, isso é um pacote muito interno.

Basicamente, ele pega, ele pega ali, o que você está fazendo, as suas variáveis, roda um light GBM, é uma estrutura machine learning para descrever o JLM. Então, tem um passinho aí que é subjetivo, né? A subjetividade está incluída ali nesse contexto. Basicamente, para elencar as variáveis, então, utilizando essas variáveis, a gente usa como, vou destacar o modelo utilizando essas variáveis, para ver se eu consigo reduzir device e já ter um comportamento inicial.

Depois eu só vou corrigindo o risco de acordo com características que eu estou montando nessas variáveis. Então, inicialmente é assim. Depois a gente passa para uma análise temporal, uma análise de estabilidade, uma análise de, aí assim, a gente chama de variável de controle, né? Essas variáveis de controle, elas fazem no tempo, para ver se o comportamento está estabilizado no tempo.

Se teve alguma mudança, alguma coisa nesse sentido. Ultimamente, a gente está usando a variável de COVID, que mesmo relacionada com o tempo, a gente usa um pré-COVID e pós-COVID. Então, para ver o que aconteceu ali na COVID, porque teve uma queda de sinistralidade, pouca movimentação, então, tem fenômenos que voltam ao patamar que estava no pós, tem fenômenos que ficam abaixo do patamar que estava na COVID.

Então, a gente utiliza essa variável para conseguir equilibrar esse comportamento. Tudo com baseline no último momento, né? No último mês, na última estrutura. E, assim, outras variáveis que a gente quer ir a controlar.

Teve variável, por exemplo, de pessoa jurídica, pessoa física. Algumas variáveis de corte de crédito. Então, tem algumas outras variáveis que a gente

utiliza como controle também. E aí, a gente entra, isso é o analista tomando a decisão ali no momento de fazer o modelo. Beleza, ele fez, ele entregou e acha que aquele modelo é ótimo. Faz uma troca entre os modeladores. Aí, a gente chama de químio. Entre os modeladores, primeiro a gente troca, aí a gente faz uma análise. Então, tem uma documentação, todas essas variáveis.

O que a pessoa que está fazendo? Ah, eu vou colocar uma feature individual para cada uma, pontuando cada nível. Ah, eu vou colocar uma curva de grau 2, uma curva de grau 3. Ah, eu vou colocar um ensemble, uma curva e uma parte de fator. Então, essas decisões você tem que explicar nesse documento.

Tem um documento inicial e a gente explica nesse documento. Coloca o gráfico lá do que está acontecendo, por que você tomou essa decisão. E aí, vem o preview em cima desse documento que você montou o modelo e ele vai, olha, não concordo com isso. Tentei a curva de grau 2 e ficou melhor ajustado. Ou aumenta o grau, porque a gente entende que essa variável passa a partir. Eu queria até apontar isso.

O seguro, ele tem uma, muito forte a questão do negócio. É muito forte. Então, tem variáveis que não podem inverter. Não pode inverter. Tem que ser decrescente, de nível decrescente. E, por exemplo, às vezes eu tô montando, aí inverte alguma coisa no preview, observa que não.

Então, a gente tem esse controle. E uma terceira camada disso, quando sai do preview, volta para o analista, a gente faz os ajustes, que aí é assim. Os ajustes também, é o preview, a decisão final é nossa. O preview, ele dá uma proposta. Ah, não, não acredito que aumentar um grau, por exemplo, por parcimônia, eu não aumentaria esse grau. Ou o aumento do grau não é significativo. Eu não entendo qual é o aumento significativo. A gente toma essa decisão. E aí, entra o manager review da última decisão, que aí são pessoas mais seniores.

Não são mais seniores, são gerentes mesmo, a galera da gestão. E entra nisso. Até o diretor também entra na rodada, que aí ele toma realmente e diz assim, não, ah não, beleza que você não concorda com o grau 3, eu entendo seu posicionamento, mas vamos aumentar, porque queremos um modelo, mesmo que com um grau maior, que acerte melhor, sabe? Que tenha uma força, alguma decisão. Então, tem essa última camada de manager review também, que é...

Entrevistadora: Assume as responsabilidades, né?

Caio: É, e assim, normalmente é conflito, eles decidem, ou tem alguns que refazem. Do tipo, beleza, eu entendi o que você quis dizer, eu entendi o que o prior review fez, só que eu vou tentar refazer de uma outra forma. Aí eles tentam de uma outra maneira, e também são mais experientes, tem outras

visões, aí tenta refazer.

Assim, às vezes dá certo, às vezes não dá, mas mesmo quando não dá certo, coloca lá escrito, ah, tentei fazer determinada coisa, porque fica como experiência pra gente, ah, beleza, então tem como testar isso. Por exemplo, acontece muito de juntar fator com curva, e já aconteceu de muito menos de ver o PDI e falar, não, é melhor juntar duas curvas do que juntar fator com curva. Então, tem essas características e essas decisões que estão um pouco diferente do que a gente faz.

Entrevistadora: Entendi, tá legal, obrigada pela explicação. Vamos, então, começar a nossa simulação, né? Eu vou passar um cenário para você, um cenário hipotético, né, e aí eu quero saber como você chegaria na solução, que você consegue planejar para esse cenário, tá? Então, qual o objetivo? É, imaginar que uma empresa de tecnologia de porte médio, tá, contrata uma outra empresa para fazer uma divulgação de vagas, ou seja, se a empresa de tecnologia tem vaga aberta, e ela quer que uma outra empresa anuncie, faça um algoritmo para conseguir anunciar essa oportunidade, que tem essas seguintes características, somente essas características que a empresa dá. É uma vaga de tecnologia para um profissional pleno, é um desenvolvedor para ocupar a vaga de um desenvolvedor ou desenvolvedora full-stack, com formação em engenharia de software, tecnologia de informação, sistemas, afins, né, estatística, enfim.

Uma experiência de cinco anos, mais ou menos, e um regime remoto, tá? Então, essa empresa de porte médio precisa ocupar essa vaga, passa esses requisitos somente, contrata sua empresa e você vai ser responsável por fazer um algoritmo de recomendação, talvez, para divulgar esse anúncio para as pessoas que têm o perfil que você vai considerar, vai fazer uma leitura, que tem o perfil que pode encaixar nessa vaga, tá? Então, a primeira, o primeiro passo seria, como você faz esse planejamento para fazer a concepção dessa solução? Em relação à definição de requisitos, né, como você define em cima dos requisitos que foram dados, e os dados utilizados, como você utiliza, como você captura, extrai e trabalha esses dados?

Caio: A definição de requisito... Perdão. Normalmente, o que eu faço muito aí, assim, é uma característica, eu acho que vem da minha formação, do pessoal que eu trabalhei, tá? Eu, normalmente, desenho aonde eu quero chegar. Então, para mim, eu estou... Encontrar a pessoa é a característica final, então, eu faria o reverso.

E delimitaria quais são as variáveis que faz sentido eu ter em relação a cada uma dessas variáveis que a gente tem como delimitador. Primeiro, eu colocaria no contexto geral e, depois, eu filtraria. Estou considerando que vai

ter tecnologia, não estou entrando nesse mérito de pesquisa e tal.

Vai ser um processo rápido, tá? Entraria com filtros mais abertos, por exemplo, filtro de cinco anos, essas coisas. Uma vaga de plenos. Então, assim, eu entraria com filtro de júnior a sênior, uma... Sem experiência, sabe? E esses maiores, e daí, teria uma base de dados com esse tipo de informação.

Depois, eu começaria... Aí, sim, farei o processo mais analítico de filtrar, de chegar no nome ou umas características de pessoas para isso. Os dados seriam essas duas variáveis, assim, bem específicas. Como os dados foram coletados.

Eu acho que poderia ser um app scrap, alguma coisa nesse sentido. Eu não sei como... Eu até estaria entrando em outro mérito. Como funciona o LinkedIn.

Eu daria alguma estrutura nesse sentido de conexão, ou se dá para gerar uma API, alguma coisa nesse sentido, que eu conseguiria estruturar num data frame. Seria essa conexão inicial. E o processamento mais... Aí seria essa questão de filtros, no momento da coleta, o processamento mais nesse sentido.

Primeiro aberto e depois eu ia filtrando para chegar no nome. Eu não sei se eu colocaria um algoritmo de recomendação, por conta do limitante da vaga. Eu vejo que ela está restringindo... Ela tem muita coisa restrita.

E eu desenvolver um algoritmo para isso. Eu não sei que eu tenha um dado de recomendação. Eu sei que eu tenho um histórico muito forte.

Se eu tiver um histórico com esse tipo de vaga, tem determinadas características, talvez eu consiga com tais perfis, aí talvez um algoritmo de recomendação entraria. Mas se eu não tiver, vai partir do zero.

Entrevistadora: Tá bom. E aí como é que você desenvolveria esse algoritmo? Você já começou a falar um pouco. Qual método você escolheria? E se consegue vislumbrar algum desafio técnico que talvez até impacte na escolha do seu método? E, por fim, a integração e implementação que você faria com esse sistema de divulgação de vagas?

Caio: Eu vou considerar que eu tenho um histórico, que eu acho que fica mais fácil de trabalhar do que se partir do zero. Vamos dizer que tem um histórico... Eu acho que nem precisa.

Eu acho que a gente pode montar uma base histórica. Eu montaria uma base histórica de pessoas que estão trabalhando nessa área, com caracteres bem similares, e marcaria ela com uma flag de um ou zero. E pessoas que estão na área ali delimitada de full stack e outras coisas, e que não trabalham com as características que estão pedindo na base.

E aí, como eu não preciso de uma força de negócio, força de característica, eu preciso mais da classificação, eu partiria direto para uma técnica mais de

inteligência computacional do que uma técnica mais clássica de estatística. Então, eu iria mais para uma técnica de classificação mesmo, machinando, uma árvore, alguma coisa nesse sentido. Se eu continuasse ainda com determinadas... Com uma dificuldade de tomar decisão em relação à probabilidade, com probabilidades muito altas para determinadas características, eu iria para um outro tipo de algoritmo.

Eu já entraria no algoritmo mais profundo, ou o algoritmo de rede neural, já que existe tanta questão de negócio. Teria que ver muito a característica dos dados e como está sendo feito esse data frame, essa estrutura para eu conseguir decidir qual algoritmo utilizar em si. Mas eu normalmente tomo decisão em cima desses steps.

O neural, como é muito forte, ela é muito agressiva, eu ficaria mais no final. Seria uma das últimas possibilidades dentro desse cenário de igualdade entre as decisões. O desafio técnico é entender o processamento, tecnologia. Isso aí é um gargalo enorme nas empresas e computadores que suportam esse tipo de algoritmo. Muita dificuldade, às vezes, de processar. Os computadores mais fortes, às vezes, demoram consideravelmente.

Então, isso seria um desafio. Essa minha estrutura de escolha tem dois motivos. O primeiro motivo é que eu acho que quanto mais profundo, mais dificuldade de a gente conseguir entender. Se a gente não quer entrar nessa parte de negócio, eu já jogo um pouco mais para a ponta. Essa questão da dificuldade técnica. E a segunda dificuldade é o processamento. Uma rede neural, às vezes, dura bem mais do que o computador. Então, os desafios técnicos são esses.

Integração e implementação. Eu acho que seria uma lista com candidatos de melhores ranqueados, eu vou colocar assim, na classificação que eles tenham. Uma classificação mais compatível com a vaga para poder fazer essa integração. A integração seria uma divulgação para a equipe.

Eu não vejo o algoritmo tomando decisão. Eu acho que precisaria passar para uma equipe humana, vamos colocar assim. A não ser que seja uma outra estrutura, aí já é uma reação negativa.

Entrevistadora: Mas você disse que teria que passar para uma pessoa, mas na própria seleção não é a seleção final. Não é tirar a entrevista do RH. Não teria isso. É só levar até a pessoa. Fazer com que a pessoa, na verdade, se candidate à vaga. Ela vai chegar nela e tal. Então, para fazer essa primeira seleção, você acha que teria que ter um aval humano? Uma escolha humana?

Caio: Não. Assim não. Eu estava pensando que a pessoa humana iria contactar. Nesse sentido não. Eu acho que pode gerar um ranking. Nesse contexto, talvez poderia fazer... Lá no final, a gente pode ter feito um cluster

ou alguma coisa nesse sentido. Ou um score da probabilidade associada. Para poder trazer um grupo que ficaria exposto a essa vaga. E aí entraria em contato, entre aspas, pelo organizador de demonstração. Apresentar a vaga no LinkedIn, receber recomendação, fazer patrocinada.

Entrevistadora: E aí, no final, são duas coisas que estão conectadas aqui. Em relação aos testes e a validação. Quais tipos de testes você faria? Métricas de sucesso, porque você está vendendo um trabalho. A empresa te contratou para fazer isso. Você tem que dar um retorno ali, através de métricas.

E até para você também considerar feedbacks e ter ações. Considerar feedbacks humanos, não humanos, para entender se está alcançando objetivos, se está sendo efetivo.

Caio: Tipos de testes. Eu sempre penso que isso precisa ser testado. Não existe 40, 40, 20, 80, 20. Para mim, isso precisa ser testado. Porque a distribuição altera, altera comportamento de variável, altera um monte de coisa. Então, eu acho que isso precisaria ser testado. E eu colocaria algumas 40, 40, 20 aí, de teste, treino, teste e out of sample.

E a quantidade eu testaria em relação ao percentual, em relação a características dos dados. Eu faria teste mesmo de colocar isso para processar no modelo básico, aí eu não faria tecnologia, não faria nada disso. Colocaria no contexto básico, para eu entender para onde que o modelo está carregando o meu teste, o meu treino, para onde que seria a estimação do meu teste, para onde que estaria o out of sample.

A gente chama isso de experimentação. A gente faz alguns tipos de experimentação em relação a isso. Não só isso, como diversas outras coisas.

Eu acho que seria em relação à modelagem, seria isso. O que eu faria mais de teste, só que aí seria um processo anterior, um processo de análise exploratória. Eu entraria em um monte de PSI, entraria no comogorógico e esmirno para ver se está desviando, entraria em diversos outros tipos de análise.

Esses processos, normalmente, são anteriores. Eu faria isso antes de entrar com isso na modelagem. Correlação, para ver se entra alguma coisa multicorrelacionada.

Se tiver variáveis multicorrelacionadas, separar para fazer um componente principal, alguma coisa nesse sentido, para tentar reduzir essa dimensão.

Entrevistadora: Entendi.

Caio: Os testes seriam mais ou menos nesse sentido.

Entrevistadora: E aí a última parte, desculpa, você quer colocar, falar de feedback, né? Métrica de feedback.

Caio: A métrica seria a curácea, basicamente, e a gente entraria no F-score ou alguma coisa nesse sentido. Mas eu abriria isso por variável. As

variáveis que estão no processo, ele precisaria ser aberto. Então, a estimação, se eu não me engano, precisa ser um profissional. Eu não lembro mais as características da vaga, mas eu abriria por essas para entender realmente o estimado, nesse sentido.

E feedback, eu acho que precisa ser contínuo. Contínuo no sentido de, ou toda semana, ou de 15 dias, as pessoas externas estarem incluídas, para poder, primeiro, entender o que está acontecendo e, segundo, para poder palpitar. Porque, às vezes, as pessoas que estão fora não conseguem entender mesmo o que está acontecendo, ou estão cansados do que estão entendendo ali, não conseguem ver se tem alguma coisa que está desviando.

Entrevistadora: Essa parte que você falou, que ia abrindo por variável, é justamente para entender se está alcançando de acordo com aqueles requisitos lá do início, né? E o que está alcançando e o quanto está alcançando, para você ir ajustando e chegando em um equilíbrio ali. Seria isso?

Caio: Porque, às vezes, eu preciso... Na estatística clássica, eu colocaria interação, tentaria forçar alguma coisa nesse sentido. No machine learning, eu colocaria alguma coisa mais profunda, ou tentaria fazer a criação, aumentar o número de árvores, ou entender como é que está a conversão dos hiperparâmetros, para ver se a convergência está indo para um lugar... Ou mudar a convergência dela.

Tomaria essas decisões para entender para onde está indo. Porque, às vezes, na métrica global, não está maneiro, está direitinho. Quando a gente começa a abrir, aí começa a desviar.

Isso acontece muito numa...

Entrevistadora: Entendi. Você começa a detalhar, o negócio está desequilibrado. E o final, assim, é um pouco já que você vem falando, mas, assim, a questão da efetividade mesmo, para você avaliar o algoritmo em relação ao filtro de candidatos que foi feito.

E, se você considera que o seu algoritmo, dependendo do jeito que você modelar, do jeito que você fizer esses ajustes nas features, qual seria o impacto na contratação, né?

Caio: Eu acho que a questão da efetividade é essa questão de análise do modelo em si, de estatísticas do modelo, de acurácia, desse sentido global e mais restrito. É importante não overfitar, né? Para não deixar só acertando naquelas características, quando a gente olha para o resto do mundo, tem que ter esse cuidado. Mas eu acho que é nesse sentido, em questão de efetividade, chegar num grupo de pessoas que têm os esquilos desejados e que sejam plausíveis para a contratação.

E acho que pode gerar um impacto porque seleciona. É mais fácil do

que ter uma pessoa só ali, uma pessoa que eu digo aí, sim, um analista de RH olhando, procurando a pessoa, tentando fazer filtros sem nenhum reconhecimento de padrão. Eu acho que, às vezes, o padrão é necessário, nesse sentido de achar a pessoa certa em determinadas características para ter uma posição.

Entrevistadora: Tá legal. A gente acaba aqui, né? Queria saber se você tem mais alguma observação, alguma coisa que você achou que faltou ter aqui, que você geralmente na sua rotina de modelagem usa, como seria? Teria alguma observação diferente? Fique à vontade.

Caio: Acho que de rotina, eu acho que o DJ é uma coisa muito forte, muito forte. E o DJ, eu acho que ele manda, principalmente no machine learning. Então, quanto mais forte o seu processo de DJ, mais forte é o seu modelo. Porque é ele que vai te dar o norte, é ele que vai guiar as suas decisões, é ele que vai ter um direcionamento específico para onde você vai ir.

E o Future Engineering também é importante, já no segundo passo. Faz as experimentações em Future Engineering de, beleza, essas três variáveis aqui que o meu DJ disse que está relacionado. Eu vou criar duas dimensões, eu vou fazer um outro processo, outras características.

E eu acho que também uma grande dificuldade, lá para o final, é a implementação disso. Quando a gente fala de web é mais fácil, mas quando a gente fala de sistema da empresa, ou, por exemplo, se fosse necessário implementar um software de um sistema dessa empresa específica que contratou, que queria, aí começa. Porque aí você tem restrição de variável, você tem restrição de performance, tem restrição de processamento.

Então, eu acho que, eu até brinco com isso, com o pessoal, que às vezes a gente tem muita metodologia, tem muita maneira de se fazer, tem várias decisões, várias formas de tomar decisão, mas que a gente não tem, ou não tem processamento, ou não tem robustez de dados, ou não tem arquitetura, ou não tem tecnologia. Eu acho que a gente está avançando, já avançou muito, mas a gente ainda está avançando nesse sentido.

Entrevistadora: Legal. Interessante. Muito obrigada, Caio. Eu vou colocar aqui o link do formulário, ver se você consegue pegar.

Aí, quando você puder responder, é só para ajudar também eu fazer um cruzamento de dados, né, na hora de analisar. Mas, assim, foi muito legal tudo que você falou, super detalhista, vai me ajudar bastante, viu?

Caio: Fala demais.

Entrevistadora: Não, é detalhista, é diferente, não é que fala demais. Você conseguiu. É exatamente isso que eu quero analisar, sabe? Qual é a estrutura de pensamento, o encadeamento das ideias, e foi muito bom.

Caio: Eu te agradeço.

Entrevistadora: Obrigada a você.

Caio: Já vou responder, ah, por agora, porque senão eu vou acabar perdendo.

Entrevistadora: Tá bom. Tá ótimo. Obrigada. Bom fim de semana para você.

Caio: Tchau, tchau.

B.4

Davi

Entrevistadora: Não, claro, tranquilo. Mas isso que eu ia falar também, assim, não precisa entrar em nada sigiloso, nem é o caso, não precisa mesmo. Mas se você quiser também, assim, entender que é você, o seu CPF que eu estou entrevistando e nada vai...

Davi: Tá, show.

Entrevistadora: Esquece que eu sou da Deloitte. Vou compartilhar aqui com você a tela, porque facilita para a gente ir se norteando aqui com as etapas, tá? Mas deixa eu te dar uma contextualização. É uma entrevista semi-estruturada que eu vou fazer e aí a ideia, a gente tem alguns tópicos, tem uma situação que eu vou passar para você com alguns requisitos e eu gostaria que você fosse me dizendo qual a solução que você conseguiria desenhar para essa situação, para esse cenário que eu vou te passar.

E a minha intenção não é exatamente entender, fazer a análise da parte técnica, das suas decisões técnicas, mas sim como você vai organizar, né, seu mapa mental em cima disso e seus direcionamentos, tá bom? Se você quiser, a hora que você quiser parar para perguntar também, peraí que meu mouse deu tilt aqui, cadê você? E também, se quiser compartilhar qualquer coisa também, porque a ideia é você, a gente fala, chama no design, esse método chama Directed Storytelling e aí é exatamente isso, você ir falando qual ação você teria, tá? Aí primeiro eu queria saber, fazer uma contextualização, se você puder contar brevemente a sua interação, seu papel profissional ou acadêmico que você já teve em relação a desenvolver algoritmos, né? E qual a sua experiência, né? E o seu papel hoje também, em o que você faz profissionalmente.

Davi: Então, hoje, meu papel hoje, eu voltei a ser engenheiro, né? Saí de carro de giro, agora eu tô mais técnico, hoje eu sou especialista, sei lá, e hoje praticamente eu ajudo os times nas suas entregas e a relação com o algoritmo é sempre, dependendo do problema que é proposto, usar o algoritmo como uma ferramenta para gerar algum tipo de otimização, né? Então, eu que

não crio nada ultimamente, não tenho criado, nunca criei, não tenho nenhuma patente sobre algoritmo, nunca fiz nada, mas, volte bem, o algoritmo é como se fosse uma caixa de ferramentas, os algoritmos, né? Seria uma caixa de ferramentas que eu posso usar no meu dia a dia para resolver algum problema de maneira otimizada. Isso tem algoritmos tanto para compressão de vídeos, digamos assim, para multimídia, algoritmos para LLM, que é a nova onda, AI, etc, algoritmos de deep learning, algoritmos também para clusterização, enfim, depende do que pede o problema.

E assim, na minha vida profissional acadêmica, eu tive que utilizar muitos algoritmos, né? Desde quando eu trabalhava com sistemas embarcados para TV e etc, para compressão de vídeo, otimização de algumas coisas para determinado dispositivo, do mestrado também trabalhei com algoritmos voltados para paralelismo, então, utilizando algumas técnicas para rodar em cima de hardware específico como GPU e etc, enfim, um pouco eu trabalhei.

Entrevistadora: Legal, bom, só para ter uma ideia mesmo assim da parte técnica, tá? Eu vou passar um cenário para você com alguns requisitos. Imagina que uma empresa de tecnologia de porte médio contrata uma empresa em que você trabalha, que faz divulgação de vagas, ou seja, faz algoritmos para rodar anúncios de vaga.

E a oportunidade é a seguinte, eles passaram esses requisitos, é uma vaga de tecnologia para profissional pleno, com skills desejáveis, né? Que contemplem papel de desenvolvedor, desenvolvedora full stack, com formação em engenharia de software, tecnologia da informação, sistemas e afins, e uma experiência aproximada de cinco anos em regime remoto, tá? Então, esse é o panorama. Só para explicar um pouquinho mais, imagina esses anúncios que aparecem aí no Facebook, no Instagram, enfim, eles já fizeram algum, o algoritmo já direcionou para aparecer para você especificamente, né? Então, a ideia é justamente isso, capturar pessoas que tenham o perfil que vão dar méritos essa oportunidade, que podem ser que queiram se candidatar, e aí aparecerem candidatos lá para eles serem, para eles entrevistarem, que tenha mais a ver com aquilo, já tenha passado por um filtro, vamos dizer assim, tá? Aí a primeira parte é em relação a planejamento e concepção, né? Em relação às definições de requisitos, né? Como você define ou refina, vamos dizer assim, né? Os requisitos para esse algoritmo, e quais seriam os dados utilizados, que tipo de dado que você imagina que pode ajudar a responder aqueles requisitos. Vou voltar lá para você ver quais são.

Davi: Tá, acho que como definição, é a primeira coisa que a gente, eu fatiaria o problema em algumas partes, né? A primeira parte seria, isso é um problema de, vamos considerar assim, de, até pensei aqui na palavra, mas

esqueci, é de similaridade, né? Eu tenho que saber qual, digamos, qual perfil, ou então qual entidade que eu tenho no meu banco de dados, alguma coisa assim, tem mais similaridade com a proposta que eu tenho aqui dentro do texto. Então a primeira coisa que eu faria era tentar encontrar, digamos assim, determinar quais palavras-chave, que seria até a resposta do segundo, mas quais palavras-chave que eu tenho que ter diante do meu banco de dados para poder criar esse tipo de similaridade. Então, por exemplo, eu definiria esse tipo de coisa, primeiro, as palavras-chave, tipo, quais minhas habilidades, qual o meu perfil, se eu sou desenvolvedor Fustec, se eu sou outro tipo de desenvolvedor, conseguir ter tempo de minha formação, quanto tempo eu tenho, experiência também, e qual a minha preferência de regime, por exemplo, vamos colocar essas aí.

Então, essa é a primeira coisa. A segunda é como eu posso, digamos, pegar esses dados e conseguir, de alguma forma, taguear isso, digamos assim, como eu posso pegar o dado que eu tenho, o input seria, definir o input, o input seria o perfil da vaga, tipo esse texto aí que você passou. Então, o que eu definiria? Cara, com esse input, o que eu tenho que ter como saída para poder gerar a minha similaridade? Então, eu teria que fazer um algoritmo para extração de pontos-chave daquela minha vaga.

Então, eu passaria por isso, consideraria uma outra parte que seria, digamos, eu chamo de, vamos chamar de injeção de dados, vamos chamar assim, Ingestão. Ingestão, isso, ingestão, perdão. Ingestão de dados para poder, digamos, gerar palavras-chave para poder ter determinada minha similaridade.

E tem uma outra parte que é como o meu banco de dados eu geraria uma saída melhor para poder calcular essa similaridade. Nesse contexto todo, eu definiria assim, ah, qual o melhor algoritmo que eu posso usar nisso? Existem várias formas, eu posso, digamos, ranquear isso por palavras-chave, por quantidade de, tipo, usar palavras, formação, tempo de formação, se for maior ou igual, eu consigo calcular isso, taguear em, digamos, em tabelas e bancos de dados e poder fazer uma similaridade, se está de acordo com o valor-chave, se estou de acordo com o best skills, eu consigo fazer um cálculo, digamos, uma inferência ali para determinar qual é a porcentagem em função dessas palavras-chaves. Eu poderia fazer isso ou poderia apelar para uma coisa um pouco mais complexa, do tipo, eu poderia, pelas palavras, digamos assim, eu tenho vários perfis de desenvolvedor de forma escrita, digamos assim, é um texto corrido, tipo Davi, Davi Prata, desenvolvedor pleno, com cinco anos de experiência, gosta do trabalho remoto, tem formação em engenharia de sódio, fazer um texto corrido, e em função disso, quando eu jogar a minha pergunta, eu posso calcular, digamos assim, eu poderia, digamos, calcular os

pesos daquelas palavras e ver qual, digamos, qual candidato ali do meu banco de dados tem maior similaridade com o texto completo.

Então, tem duas formas de fazer, mas, assim, a grosso modo, eu escolheria uma maneira para poder resolver e depois eu resolveria, digamos assim, depois eu aplicaria isso, e sobre os dados necessários que eu vejo, a gente, digamos, você tem que ter dados dos possíveis candidatos, de alguma forma, então, você tem que, vou dar um exemplo do LinkedIn, o LinkedIn, você coloca a sua informação, você faz o seu resumo, o seu, desculpa, o seu currículo, e aí, com o seu currículo, eles conseguem ali mapear e taggear o que você tem de habilidade e, com isso, quando gerar vaga, gera uma similaridade, então, eu tenho que, digamos, os currículos dos candidatos, com os dados que são propostos para ver similaridade, como formação, visibilidade, hard skills, talvez soft skills, não sei, depende do que você quer procurar naquilo, a quantidade de experiências e também com a preferência do Regime, às vezes, você consegue preferência de Regime, ou preferência de outra coisa, através de outra função dentro da sua aplicação, digamos assim, o LinkedIn, ele não fala a sua preferência, não coloca a sua preferência no site, às vezes, você faz buscas por emprego daquela preferência que se torna, digamos, a preferência desse cara é para pagar remota, então, vamos tentar jogar a vaga remota para ele, então, depende de como os dados são esses que eu falei, e para você pegar esses dados, depende de como sua aplicação vai interagir, isso é mais uma questão de UX mesmo, de uso da aplicação em si, da rede social, do aplicativo, não importa o que for, para poder pegar dados dos usuários daquela rede, para, digamos, ter esse cálculo de similaridade.

Entrevistadora: É quase que ir cruzando a informação com a usabilidade ali.

Davi: É, só que, eu estou falando assim, pela experiência que eu tenho, normalmente, o que os caras fazem hoje? O que eles fazem? Eles fazem você responder coisas rápidas, porque hoje todo mundo está incentivado a ver vídeo curto, tudo curto, até os forms são curtos hoje, então, o direcionamento é que você entre o mais rápido possível na plataforma com poucos dados, e durante o seu uso, você vai capturando outras informações que sejam importantes para poder fazer esse tagging, no caso de Vava, faça isso, talvez o cara no aplicativo de relacionamento, dependendo das escolhas que você faz, dos perfis que você utiliza para, ah, eu gosto de mulheres noivas, sei lá, o cara percebe que o perfil é de mulheres noivas, então, aquele tipo de foto também aparece essa preferência para ele, entendeu? E assim vai, então, de viagem, organização de viagem, foi feito dessa forma hoje, então, você consegue pegar os dados de forma direta e também de forma indireta, durante o uso da aplicação.

Entrevistadora: Tá, legal. E aí, assim, aqui a gente passa para a segunda parte, que eu acho que você já falou alguma coisa em relação à escolha de métodos, e aí, queria até pensar, assim, em relação, já pensando na parte do método em si, como você falou, tem dados diretos e indiretos, você daria algum direcionamento específico para fazer essa equação, essa regra de negócio, pegar os dados diretos e indiretos ali para contemplar esses requisitos específicos, lembrando que, assim, vai selecionar um público que vai receber aquela oportunidade, né?

Davi: Então, depende de como a aplicação vai navegar, isso aí não tem muito uma determinação de como o algoritmo funciona, você tem que coletar os dados, digamos assim, você poderia fazer um form gigante e pegar todos os dados possíveis de uma vez só, né? O motivo do cara fazer, utilizar uma forma pequena, para o cara poder usar a aplicação, digamos, viciado dela, e pular os dados indiretamente, é uma forma de reter o usuário, é uma coisa mais voltada, como eu falei, para a UX. Sobre a escolha do método, assim, para esse tipo de questão, eu utilizaria o método de NLP, porque você trabalhou com questões textuais, né? Você poderia taggear, só que qual o problema de taggear? Você deixa, digamos assim, você pode gerar uma, quando você cria tags, quando você cria, assim, você tem que criar pacientes, mas, tipo, você pode fazer o seu perfil totalmente voltado para aquela coisa, porque você já sabe que o algoritmo funciona de uma forma que já extrai palavras-chave, digamos assim.

Você também pode escrever um texto com palavras-chave, só que se você tornar ele mais, digamos, acessível a todo mundo, porque você pode fazer o seguinte, vamos dizer que eu escrevi o meu perfil, mesmo com pontos, etc. Dali, com aquelas informações, eu gero um texto, digamos, corrido, padrão para todo mundo, de como é o perfil da pessoa. Tipo, Larissa trabalha com UX, enfim, trabalha com desenvolvimento, há mais de cinco anos, formada pela UFRJ, mestrado, doutorado, hoje trabalha tal coisa.

O que acontece? Você consegue fazer uma coisa, como é que funciona?

Quando você utiliza esses algoritmos voltados para a mastilândria, para você conseguir calcular, por exemplo, a ponderação, o peso daquela palavra, qual próxima palavra eu posso lançar, qual próxima palavra eu posso lançar. Então, você consegue fazer, utilizar também, você pode, digamos, ter algoritmos, você pode processar isso e gerar, tipo, esses pesos, que são chamados em média, e você pode, digamos, calcular a proximidade de um texto para o outro, digamos assim, um para um. A OpenAI funciona assim, né? Digamos, um chat de EPT.

Ah, eu vou escrever alguma coisa. O que ele faz? Não é que ele saiba tudo, ele faz o cálculo do próximo, qual melhor palavra que pode ser resultado para aquilo. Só que, para ele fazer isso, ele utiliza esses embeds.

O que ele faz? Ele busca ali, através da semântica dele, assim, cara, essa palavra aqui, essa frase aqui, considerando todas as palavras que são, digamos, são tokens. Estou generalizando, tá? Não é bem isso aí não, tá? Mas, que são tokens, né? Porque token pode ser uma palavra, pode ser, uma palavra pode ser dois tokens. Uma palavra, tipo, o verbo do gerúndio, tá ligado? São dois tokens.

Porque você quebra, enfim, é uma questão de... Mas vamos colocar assim, ali eu tenho um grupo de tokens ali. Eu posso calcular, digamos, qual o grau de proximidade, de singularidade, né? Porque você pega as ultrapartes palavras, transforma em vetores, e você consegue calcular, digamos, a similaridade, a proximidade de uma frase para a outra. Então, muitas vezes você só precisa taggear.

Tipo, vou dar um exemplo assim, eu quero taggear a palavra sorvete, né? Só que aí, vamos dizer que o candidato escreveu lá, sei lá, gelato. Eu gosto de gelato. Muitas vezes vai ter uma similaridade.

Se eu fizer por tag, eu não vou pegar essa similaridade, porque eu vou taggear quando o cara falar sorvete, alguma coisa assim, ali na demonstração. Se eu utilizar a similaridade, eu sei que gelado tá muito próximo de sorvete, então eu posso gerar um grau de similaridade de uma frase para a outra. Então é possível, digamos, gerar um filtro muito melhor.

Então, pra esse caso, dessa solução, eu usaria um algoritmo de LLP, utilizaria, aplicaria lá um algoritmo para você poder calcular a similaridade pelos invés, por exemplo. Cara, pode continuar aí, se quiser.

Entrevistadora: Não, eu acho que assim, o desafio, você falou já de algumas questões, inclusive essa questão do porquê que você escolheu, escolheria por similaridade.

E vou passar para a terceira, na verdade, que é em relação a testes e validação. Se você consegue pensar em algum teste que você realizaria, e métricas de sucesso também, porque lembrando que você vai vender essa solução. É um produto que você está fazendo que vai vender para o contratante.

E, por fim, feedbacks e iterações. Se você considera, acha importante, capturar feedbacks humanos e não-humanos para fazer a melhoria no algoritmo.

Davi: Sim.

Então, a melhor forma de fazer os testes é você injetando mais dados. E depois, como é que eu faria o teste? Eu teria dados, vamos chamar de dados bloc, dados meus, criados por mim, com grande quantidade, e geraria uma entrada e saída. A questão toda é que eu tenho que bater todas as minhas entradas com as minhas saídas para poder passar no teste.

Eu quero saber qual é o candidato de um grupo que eu tenho que melhor se encaixa para esse tipo de vaga aqui. E só pode ser um. E o meu algoritmo tem que bater, não importa quantas vezes eu mudar, mas o meu algoritmo tem que bater com o candidato A, por exemplo.

Se não bater com A, tem um erro. Então, eu faria esse tipo de teste para poder resolver. Então, nesse caso, estou falando só da questão do algoritmo, nem falo do sistema como um todo.

O algoritmo em si, eu pensaria em fazer apenas testes, digamos, de unidade. Eu criaria mocs de possíveis valores para input, teria saídas, eu determinaria outputs que deveriam ser iguais ou próximos àquele output, digamos assim. Tipo, qual o candidato melhor para se encaixar nessa vaga, qual o candidato melhor para se encaixar.

Quantos candidatos estão de acordo com esse perfil aqui e aí listar aqueles candidatos. Então, faria todos os testes relacionados a isso. As métricas de sucesso que eu utilizaria, cara, uma coisa que eu não falei é a questão de velocidade.

Então, eu acho que a gente pode considerar como uma métrica de sucesso, performance desse algoritmo, quanto tempo ele pode dar uma resposta e também a taxa de acerto dele. Tipo, quão assertivo ele pode gerar a sua solução. E feedbacks e interações, eu acho que isso aí seriam com testes, digamos, como se fossem testes manuais mesmo, digamos, com pessoas, com possíveis clientes.

Então, vamos dizer que eu estou fazendo uma validação nesse caso. Eu fiz uma POC. Aí eu pego minha POC e coloco para 20 usuários diferentes com aqueles dados.

Eles vão lá, fazem as inscrições e entram na minha rede, por exemplo. Aí eu crio as vagas e eu mando para eles essas vagas. Então, a minha validação seria feita com uma quantidade de usuários.

Assim, 20 eu acho muito pouco, talvez 50 usuários, 100 usuários. Tentaria mandar o máximo que eu pudesse, porque cada vez mais dados executados, melhor a validação. É mais difícil para analisar manualmente, porque tu vai ter que ver um por um para ver se está batendo alguma coisa, se tu vê que tem alguma coisa que está atrapalhando.

Mas eu faria dessa forma. Pegaria uma quantidade de pessoas para poder utilizar o algoritmo mesmo, para validar. Porque o teste unitário fica um pouco viciado em alguns aspectos.

Quando você tem a etapa de você fazer algo, vamos chamar assim, de laboratório, uma coisa da... In-house, vamos chamar assim. Você faz uma coisa ali, controlado, é diferente quando você coloca isso para pessoas, para o ambiente não controlado. Tu não sabe o que a pessoa pode escrever, então

tu começa a ter algumas análises do tipo, algumas entradas que não estavam sendo pensadas, que podem confundir.

Enfim, tem uma série de coisas que você pode pegar nesse meio, nesse teste de validação com os usuários. Seria uma etapa mais posterior, depois que está rodando bem, os testes provocados ficaram bem, eles têm uma boa performance, digamos. Performance por quê? Também você não pode gastar, digamos, uma hora para o usuário poder calcular esse tipo de proximidade.

Então você tem que conseguir fazer algumas otimizações para poder acertar isso. Também você tem que fazer algumas melhorias, tipo, resolver algumas melhorias para você poder evitar falsos positivos, coisas do tipo. Então, com as métricas você consegue deduzir isso, ter uma direção e os feedbacks, a parte de feedback seria teste para terceiro mesmo, também usando métrica.

Mesma coisa, performance, que eu acho que é válido, e também assertividade.

Entrevistadora: Legal. E aí, em relação à efetividade e ao impacto, porque considerando que eles vão utilizar o algoritmo para fazer esse filtro, fazer o recorte já de pessoas que têm mais afinidade, com a vaga, então, de fato, vão selecionar um grupo, o que isso pode ser efetivo, o quanto isso pode ser efetivo e o quanto isso pode impactar na contratação e na sociedade como um todo, porque já é um recorte que está sendo feito.

Davi: Sim. Então, tem uma coisa, a questão de efetividade, eu entendo que você faz o recorte de candidatos que têm aquelas skills desejadas, digamos assim, dependendo de como o seu algoritmo pode estar sendo voltado para uma direção específica. Eu entendo.

O que eu acho que você pode avaliar é ver se, de fato, aqueles candidatos ali estão, de fato, contemplados para aqueles tipos de vagas. E a melhor forma de você ver, eu acho que vale a pena você fazer, é como se você fizesse a pergunta, ver a possibilidade daquele candidato se candidatar à vaga. Vou dar um exemplo para você.

Se no meu teste o candidato recebeu aquelas oportunidades, os possíveis candidatos, e se candidataram àquelas vagas, então, sei que está sendo efetivo o uso do meu algoritmo. Uma visão que eu tenho. Eu acho que eu veria assim como efetividade.

O impacto da contratação, como algo que pode gerar um impacto criativo, eu acho que ele pode se tornar um pouco imparcial em alguns momentos, digamos assim. Mas eu acho que esses algoritmos normalmente são usados em uma fase pré-candidatura, vamos dizer assim, seria para fazer um furio, para você reduzir o furio. Então, eu vejo um impacto... Eu vejo

uma melhoria no processo de contratação, porque você dá uma entrevista à candidata que não tem nada a ver com aquele perfil.

Vou dar um exemplo. Às vezes eu recebo vagas que não têm nada a ver com o perfil. É bem grande, é assim.

É bizarro. Todo dia eu recebo uma vaga que não tem nada a ver com o meu perfil. Nada a ver, nada assim.

Vem de headhunter. O algoritmo desse perfil, com certeza, simplifica o trabalho desse headhunter ou de qualquer outra pessoa, no início do processo de hunter, de caça ao candidato, ele consegue simplificar a ponto de conseguir tentar encontrar, de fato, o candidato mais indicado. E para o candidato é mais interessante, porque ele chega nas vagas que são realmente, de fato, de interesse dele.

Então, eu vejo que o impacto para a contratação é mais nisso. Então, você não tem um processo seletivo longo com muitas pessoas, e rola muito na área de TI. Tipo, abri a vaga e não contratei ninguém.

Sempre falam isso. Eu não acho que o problema seja relacionado com que não tem ninguém do mercado. Na verdade, tem muita gente do mercado.

Eu acho que as vagas não estão chegando na pessoa certa. Nas pessoas certas, digamos assim. Então, essa é a minha visão.

Só que, de novo, o algoritmo sempre vai ter que chegar e fazer um fine-tuning dele. Vai ter que ajustar ele. Não importa.

Não importa o algoritmo que você usar. Vai ter que fazer um fine-tuning. Porque você pode gerar um impacto, em algum momento, até negativo.

Como assim? Se você não tornar ele, se você não fizer algo que torne mais imparcial, pegando realmente, de fato, aquelas esquinas dos conhecimentos, você pode gerar, digamos, um caminho do tipo, só vou escolher candidatos formados em universidades de São Paulo. Uma coisa maluca, no exemplo. E você tira, digamos, o eixo... Você perde talentos fora, entendeu? Que não pertencem a São Paulo, por exemplo.

Então, você tem que fazer de alguma forma que isso torne mais... É um trabalho que é bem interessante, que eu acho que é, tipo, tornar mais... Até popular, não é? Mais acessível, não é? Para qualquer tipo de candidato e verem de igual para igual, só em relação à habilidade. Porque tem muita gente que... Tem muita gente, não importa de onde é, não importa o que faz, etc., que é muito habilidoso. Às vezes não se expressa bem, ou então, às vezes, não tem um bom currículo, não tem nada disso.

Eu acho que a diferença vai ser isso. Vai ser como a plataforma pega os dados de uma maneira igual para todos os candidatos, igual, igual, tem

que tentar ser igual, digamos assim. Na plataforma, não tem sexo, não tem diferença social, tem pode... O dado vai chegar igual para todo mundo.

E, dali, você consegue fazer um fine-tuning para que a vaga, para que aquela posição seja distribuída, de fato, para as pessoas que contemplam aquelas habilidades, enfim, etc.

Entrevistadora: Tá, legal. Aí, você está... Você, só interpretando um pouco assim, você diz que é possível que o algoritmo crie vieses.

Então, por isso é importante esse fine-tuning que você está falando, né?

Davi: Depende dos dados. Por exemplo, se eu pegar... Vou dar um exemplo bem maluco. Se eu pegar o currículo só dos caras da Faria Lima daqui de São Paulo, vai ter o direcionamento de só achar quem é boa e quem é feio que mora no centro de São Paulo, aqui da Faria Lima, aqui do lado.

E você só vai pegar esse perfil, digamos assim. Se você pegar só o currículo de programador formado em Universidade Pública, só vai pegar esse perfil. Então, você tem que, de alguma forma, ter uma quantidade de dados e conseguir imputar os dados de uma maneira muito transparente.

Os dados é como se não tivesse... Se eu pego dados só de uma determinada região, vai dar problema. Isso não vai funcionar. Porque você vai... Ele fica enviesado.

Não tem convite. Essa coisa que fala o quê? Ah, não existe preconceito. O racismo não existe.

Não, existe. Isso existe, sim. Você cria um viés.

Você cria uma direção. Então, qual é a melhor maneira? Você tem que tentar aumentar, pulverizar a quantidade de dados em escalas, digamos assim, de qualquer nível Brasil, em nível nacional. Então, você tem que pegar... Desde o cara que estudou lá na Universidade de Amazonas, desde a menina que estudou lá no Paraná, sei lá, juntar todo mundo.

Se esse cara se formou em Engenharia, Mecatrônica, é essa menina também. Então, a gente tem as cruzes. Os paralecidos.

Qual o problema? Qual a diferença? Conseguir ver como eles podem putar isso na plataforma que não gere tanta diferença. O que gera? Desigualdade gera. Vou dar um exemplo assim.

Vou pegar... Exemplo. Pega o currículo de um cara que veio lá da periferia, o currículo do cara que veio daqui da Faria Lima. Vou dar um exemplo.

Cara, o currículo do cara da Faria Lima é uma página só. Nossa, talentoso e tal. O cara sabe escrever, tem boa formação.

O cara, na plataforma, na hora de putar o dado dele, vai conseguir putar de uma maneira muito mais rica. E se você só usar dados daquelas pessoas ali,

digamos que seja uma solução algorítmica, quando você pega o currículo do cara que veio da periferia, a similaridade não vai bater tanto. Vamos colocar assim.

Então, você vai começar a gerar exclusão desses candidatos de alguma forma. Tipo, de quem veio desses meios. Então, a melhor forma de você fazer é, cara, a minha gama de dados tem que ser a nível, digamos, nacional.

Tem que ler um pouquinho daqui, da Faria Lima, um pouquinho lá. É o trabalho mais difícil. Aí, você vai ter que conseguir tentar entender como você pode cruzar tudo aquilo.

Aquela parte de você definir como é que você pega os dados, etc. Tá, como é que eu vou cruzar esses dados que são de outras regiões, de outros lugares do Brasil? Como as pessoas escrevem? Como as pessoas tratam isso? Você vai colocar tudo isso nesse banco de dados de uma maneira efetiva, igual para todo mundo. E a partir daí, você consegue fazer um trabalho mais efetivo e mais justo para os candidatos, digamos assim.

E você consegue gerar mais diversidade. Eu vejo desse caminho. Se você usar só... Ah, não, vou pegar os dados do meu grupo.

De exemplo, cara, eu trabalhei lá na Gabriel, aquela empresa de segurança. Cara, todos os caras. Só tinha homem, a liderança.

Só que eu usava era Rio, tá ligado? Tipo assim, era o meu perfil. Assim, sem sacanagem. O produto não consegue avançar tanto porque você não consegue descobrir outras dores de lugares mais periféricos, de outras necessidades.

Você não tem diversidade. Isso atrapalha. Então, você poderia ser isso.

Você vai fazer um algoritmo que só vai funcionar, digamos, na Faria Lima. Se você pensar, os dados da Faria Lima. Ou se você só usar os dados da periferia de uma pessoa, só vai funcionar na periferia de uma pessoa.

Por quê? Porque a similaridade sempre só vai bater ali. Então, o mais importante hoje, no algoritmo para a AI, não é o Deep Learning, não é a GPU que está rodando. O pior mesmo é você pegar os dados, que eles chamam de... Hoje, na área do elementar, eles usam o processo de Reg, Retrieve, Aumentation e Generation.

Acho que é Generate, Generation. Acho que é isso. E esse... Como é que você consegue pegar aqueles dados todos, jogar nesse Reg para que ele consiga gerar dados efetivos para que, quando você fizer uma inferência, ele traze um retorno para você.

Então, esse é o mais complexo que tenho, que eu vejo hoje, para a Machine Learning, os dados. O mais difícil é ter dados. Não é aplicar o modelo.

Isso já tem muita coisa pronta. Mas o mais complexo é você poder modelar os dados para que o algoritmo funcione. Isso que é o mais difícil.

Entrevistadora: E só a última pergunta. Você acha que seria interessante, na hora que você estiver desenvolvendo... Eu sei que, muitas vezes, vocês costumam validar com seus pares algumas decisões que vocês precisam tomar.

Você acha que seria interessante ter algum checklist, algum framework para você... Por exemplo, se tratando de questões que podem ter impacto social... E aí, nesse caso, a gente está falando de pessoas que vão ter ou não oportunidade.

Você considera que seria importante ter um checklist de coisas para se pensar na diversidade ou na inclusão na hora de fazer modelagem?

Davi: Eu acho que sim. Principalmente na modelagem de dados. Isso é o principal, na minha opinião.

Eu vejo que você tem que ter um... Eu acho que quem tem que trazer esse checklist de uma empresa seria a gerente do produto. A gerente, o PM, normalmente. Por quê? Se esse checklist... O engenheiro não vai desenvolver sem isso.

Não tem como. Ele só vai pensar em dados. Como é que eu vou pegar esses dados todos e colocar no banco de dados editorial? Vou dar um exemplo assim.

É isso que ele vai pensar. A questão de como ele vai colocar, ele vai pegar tudo e colocar. Mas se você tiver um... Por isso que eu acho que, nesse caso, tem que ter, digamos, um checklist não só para o engenheiro, mas principalmente para quem... Para o cientista de dados.

Mas isso voltado... Cara, é um projeto social. É um projeto que quer cobrir uma boa parcela da população brasileira, considerar pessoas marginalizadas, etc. Tem que ter uma pessoa especializada, na verdade.

Não é uma pessoa... Com todo respeito a quem trabalha com gerência de produto, etc. Mas tem que ser uma pessoa voltada com essas habilidades sociais. Porque senão, na verdade, você vai entregar o produto.

Normalmente, a empresa avisa nunca. Então, vamos entregar e avisa antes. Se entregou, funcionou.

Se está vendendo, é bom. Senão, pula para outro. Funciona assim.

Mas se você quer garantir assertividade, etc. Assertividade em uma questão com a popularização das pessoas, etc. Ah, eu quero conseguir que meu algoritmo alcance regiões, mas para o interior do país, etc.

Esse checklist tem que ser feito na hora da modelagem dos dados. Porque senão, você não consegue. E aí você volta àquelas coisas da validação.

Você consegue bocar melhor. A forma que escreve, etc. E é um trabalho de pesquisa.

Não é um trabalho que o cara vai definir um checklist, um framework de uma vez. Ele vai ter que pesquisar o que ele quer conseguir inferir com aqueles dados, de fato. Como ele quer ter aquela solução.

Descobrir como... Começar a ver como poderia. É um trabalho com design mesmo. É produto.

Ver como aqueles dados podem ser imputados. E ali fazer testes para ver como esses dados são imputados por diferentes partes do grupo que ele quer atingir. Quero atingir todo mundo? Vou ter que pesquisar com todo mundo.

Vou ter que buscar um cara lá do interior, um cara que mora numa comunidade, numa grande cidade. Enfim, vai ter que fazer esses tipos de pesquisa para poder tornar, digamos, a plataforma... Eu não falo preconceituoso, mas não tornar a plataforma sempre separar um grupo pequeno. Tipo, sempre ficar separando um grupo pequeno.

Sempre estimulando aquele mesmo grupo. Senão, você vai fazer uma coisa que só vai estimular o mesmo grupo. Você cria, digamos, abismos... Não tem abismo social, mas é abismo... Recortes, né? É, você cria recortes.

O algoritmo é a pior coisa, porque é uma coisa que você... Anualmente, uma pessoa consegue fazer os recortes. Você resolve de uma forma. Como? Tem uma pessoa que supervisiona, vê, avalia de novo manualmente, faz tudo.

O algoritmo é uma caixa preta. Entra A e sai B. Entra A e sai B. Então, você tem que ter essas validações para saber se o algoritmo está fazendo os recortes como você deseja. Senão, faz de novo, né? Faz de novo, digamos, melhora.

E o que vai determinar isso, hoje, para... Se você está pensando no câmbio de AI, como a gente falou, porque você pode fazer aquele tagueamento, que é o que a gente conversou, de particularidade. Mas vamos dizer que a gente quer usar a AI para resolver esse problema. A melhor forma é você criar as ferramentas de validação, como a gente falou antes, para poder testar esses casos, para saber se ele está indo para um recorte menor e se ele está conseguindo fazer o recorte que você deseja.

Entrevistadora: Entendi. Muito bom, Davi. Várias ideias boas que você me deu aqui.

Fui até anotando e vou pegar na transcrição também. Foi isso. Muito bom.

Muito obrigada. Viu? Já ajudou muito, assim. Você foi muito minucioso nas descrições e isso ajuda bastante.

Eu vou colocar aqui no chat só o formulário que eu te falei. Se possível, você... Não precisa responder agora, não. É rapidinho, tá? É só escolher ali o que você... Algumas questões demográficas.

Mas... Quando você puder responder, está ótimo. Tá bom? Obrigada. Você está no mundo aí.

Davi: Tchau, tchau. Tchau.

Entrevistadora: Beijo para você, para o Chico, para a Marília, tá bom?

Davi: Tá, vou mandar sim.

Entrevistadora: Tchau, tchau. Tchau, tchau.

B.5

Francisco

Entrevistadora: Não, não é isso.

Francisco: Só para avisar, acho que você cancelou o filtro.

Entrevistadora: Eu diminuí. Depois eu... melhorei. Tem mais resolução agora. É. Está vendo a minha cara, né?

Francisco: Dá o último goste.

Entrevistadora: Fala alguma coisa aí para eu ver se... É, foi. Foi para o português. A transcrição aqui é automática e é em inglês.

Aí você fica falando em português e ele fica tentando encaixar em inglês. Depois fica uma miscelânea ridícula. Tá, já está gravando, tá?

Francisco: Uhum, perfeito.

Entrevistadora: Tá. Eu vou compartilhar com você um PowerPoint que é só... É um roteiro, tá? Porque é uma entrevista... A gente pode chamar de semi-estruturada, mas também... No design a gente chama de directed storytelling. Que é um método que eu peço para você... Me falar, replicar o que você faria.

Tá? Numa ação específica. Então, assim... Eu costumo compartilhar a tela com... Com o texto mesmo. Com os principais direcionadores aqui para pautar a nossa conversa.

Mas se você quiser mostrar alguma coisa, alguma ferramenta, alguma... A sua tela, né? Enfim, algum modelo aí. Você fica à vontade para compartilhar, tá?

Francisco: Tá.

Entrevistadora: Então... Vou começar aqui... Pedindo para você contextualizar... Você está vendo a minha tela?

Francisco: Sim, perfeito.

Entrevistadora: Contextualizar um pouco da sua experiência profissional, no geral. E a sua experiência com algoritmos e, assim... Se você está trabalhando hoje, o que você faz? Um pouco do que você faz na sua empresa.

Francisco: Perfeito. Tá bem. Pode falar?

Entrevistadora: Uhum.

Francisco: Bem, a minha experiência com algoritmos... Ela é especialmente em algoritmos de baixo nível, né? Linguagens de máquina.

E mais conectado com o mundo da eletrônica. E tive uma vivência... Trabalhei em alguns projetos aplicando inteligência artificial. Principalmente em processamentos de imagem.

No campo de visão computacional. Atualmente trabalho um pouco mais na engenharia de modelos. E telemetria.

Em alguns casos. Ou seja, são áreas não muito especializadas. Trabalho em áreas diversas.

Mas o meu background seria basicamente esse.

Entrevistadora: Tá bom. Eu vou falar com você um cenário fictício, tá? E aí a ideia é você falar o que vier à cabeça.

Não tem problema você não ter uma experiência específica na questão de desenvolvimento de algoritmos de recomendação. Porque a ideia é entender a estruturação do seu raciocínio. Como você faz o racional.

E aí é muito importante que você vá me dizendo o que você pode pensar. Pode trazer suas dúvidas. Também dizer o que de repente está faltando para você conseguir.

Realizar alguma decisão. Mas fica bem livre. O cenário é o seguinte.

Uma empresa de tecnologia de porte médio. Contrata uma empresa de divulgação de vagas. Para anunciar uma oportunidade. Você no caso. Você trabalha nessa empresa de divulgação de vagas. Então uma empresa de tecnologia vai ser seu cliente. E eles dão só esses parâmetros. A princípio. Que é uma vaga de tecnologia para profissional pleno. Que são skills que envolvem o papel de um desenvolvedor. Ou desenvolvedora full stack. Com uma formação específica em engenharia de software. Tecnologia da informação. Sistemas e áreas de ciência e tecnologia. Uma experiência pelo menos de cinco anos. E um regime remoto. Vai ser regime remoto. Essas são as características principais. A empresa chega. Contrata a sua empresa. E pede para você fazer esse anúncio. Então esse anúncio. Não precisa ser. Desse jeito que eu vou falar. Mas você pode imaginar por exemplo. Um anúncio que vai aparecer no Facebook. Ou um anúncio que vai aparecer. Alguma coisa de LinkedIn. Você pode ir conectar com algumas dessas plataformas. Que já existem. Mas a ideia é fazer um anúncio. Para chegar nas pessoas com um perfil. Que você acha que pode encaixar. Com esses requisitos que a empresa te trouxe. Se você quiser eu volto nesse slide aqui. Mas a ideia é assim.

A gente divide mais ou menos em três etapas. Primeiro a simulação. Pensando no planejamento e a concepção. Em relação a definição de requisitos. Como você definiria. Tem os requisitos que te trouxeram. Como você elabora

esses requisitos. E como você capta esses dados. Quais seriam os dados utilizados. Para você desenvolver. Fazer a base desse algoritmo. Se quiser eu volto lá no outro slide.

Francisco: Para definição de requisitos. Não sei se aqui é o caso. De questionar os requisitos. Eu tentaria entender um pouco melhor os requisitos. Me baseando no que apareceu ali para mim. Em termos de perfil do profissional. Eu vi basicamente que são skills. Tempo de experiência. E formação. E regime de trabalho compatível. Inicialmente seria utilizar esses requisitos. Pensando de maneira que não existe nada. Que eu não vou usar nenhum outro serviço. Uma coisa que não existisse. Não existe LinkedIn. Não existe plataforma de anúncio. A gente vai fazer o algoritmo do zero. Eu poderia partir de início. Para tentar coletar dados estruturados. De profissionais. Que pudesse ser interessante. Para que eu pudesse selecionar. E abordar alguns deles. Se a gente está falando especificamente de Machine Learning. Eu posso focar nisso. Porque profissionalmente eu não começaria no Machine Learning. Para fazer algo assim. Devo manter essa linha do algoritmo? Quando a gente fala algoritmo de modelagem?

Entrevistadora: Sim. Você pode também recomendar outra tecnologia. Não tem problema. Outra solução. Imagina que realmente é uma situação. Quando você falou. Não sei se cabe questionar. Cabe questionar. Você pode levantar esses questionamentos. Claro que eu não sou a empresa. Eu não vou ter como te explicar. O que a empresa falaria. Mas a ideia é isso. Eu quero saber o que você pensa. Como você estrutura o seu raciocínio. Por exemplo. Na questão dos dados. Como você pegaria? Quais os dados que você gostaria de utilizar.

Para cumprir pelo menos esses requisitos. Contemplar os requisitos que já estão delimitados.

Francisco: Perfeito. Para melhorar o dado um pouco. Eu ia simplesmente perguntar para o contratante. Qual é o que eles estão tentando fazer. Que eles sentem que não tem o conhecimento. E qual é. Se eles pudessem dar um exemplo. De alguma coisa que eles precisam fazer. Algum caso em específico. E que eles não tem o conhecimento. Ou a capacidade para fazer. Para que eu pudesse utilizar isso. Nas minhas entrevistas. Seria o dado além desses que estão aí. Que eu precisaria. Entender qual seria o desafio dessa pessoa. Mas os dados já estão suficientes para a gente seguir.

Entrevistadora: Mas desafio como assim? Você imagina como você conseguiria isso?

Francisco? Sim. A pessoa fala. Eu tenho essa vaga aqui.

Com essas características. Mas eu gostaria de entender o seguinte. Por

que você precisa dessas características? Você já tem esse perfil na sua empresa. E está precisando de mais força de trabalho. Ou você precisa de um profissional. De um perfil novo na sua empresa. Para resolver algum problema. Que você já sabe qual é. E se você puder me dar um exemplo. De um problema. Que você não teria alguém para resolver hoje. Porque aqui eu vejo um perfil. Mas eu não vejo algo que eu possa questionar alguém.

Aqui são dados curriculares. Exatamente como você está querendo saber. Forma de pensar. Eu não vejo algo que eu pudesse utilizar. Para extrair a forma de pensamento de alguém. Que pudesse fazer o fit com a vaga. Seria a única coisa que eu colocaria mais. E também para trazer o interesse. Na pessoa que eu vou selecionar. Para que ela não sinta só. Tá bom. Eu tenho os requisitos. Para que ela também. Isso aqui pode ser interessante. Então qual seria um desafio relevante para essa vaga. A missão da vaga. Mas isso aí não é dado de entrada. Seria mais um dado de verificação para o futuro. Mas de qualquer maneira. Com esses dados a gente pode seguir. Com certeza.

Entrevistadora: Você se preocuparia mais. Em relação a dar o fit. Dar o met com a vaga. Em relação a uma coisa menos formalizada. Uma coisa mais da intenção. E da preparação subjetiva. Talvez do candidato com a vaga. Queria entender exatamente o propósito do negócio.

Francisco: O meu propósito é afunilar. Porque eu achei. Como o dado me parece um pouco genérico. Um pouco genérico demais. Eu gostaria de poder afunilar minhas buscas. Dar mais informação mais rica para o meu modelo. Como por exemplo. Full stack é tudo. Então aqui eu vou ter de todas as tecnologias. Eu não tenho uma letra. Então daqui já dá para extrair milhares de pessoas. Então eu queria tentar buscar. Já que é um tempo de experiência longo.

Eu ia querer saber. Poderia me dizer algumas tecnologias. Ou alguns desafios. Alguma área de atuação mais específica. Que eu pudesse ter uma feature mais rica. Ou seja. Enriquecer a minha feature dos meus modelos. Facilitar uns atalhos para os modelos. Em termos de experiência profissional. Ou em termos de área de atuação. Mas era mais nesse sentido. Ter um dado. Uma feature mais específica. Mais rica.

Entrevistadora: Porque. De todo modo. Quando você fala. Os dados já estão aí. Mas quais seriam esses dados?

Francisco: Por exemplo. Linguagem. Por exemplo. A de desenvolvimento iria mais fundo ainda. Quais são as. As ferramentas. Que essa pessoa provavelmente vai usar. Precisaria ter alguma experiência. E não só. O nome do cargo. Quais são as ferramentas. Que ela tem que ter algum tempo de experiência. Ou que ela tem que ter tido alguma vivência. Que aí já me direcionaria. Para

qual. Um de cem áreas diferentes. Que na área de desenvolvimento. Eu poderia ter.

E eventualmente. Eu não teria alguém de cinco anos de experiência. Mas eu teria alguém de três anos de experiência. Na ferramenta. Porque essas variáveis. Vão se cruzar lá na frente. Então. Eu poderia ter um dado. Sobre ferramenta. Tempo de experiência por ferramenta. E tentar. Melhorar o meu dado.

Entrevistadora: Então você focaria nessa parte. Da experiência profissional em si. Para não ficar. De repente. Baseado em outras coisas. Mais assim. Formais. Tipo formação.

Francisco: Sim. Eu acho que. Facilitaria. Ter a experiência. Porque são tantas ferramentas. Para especificar. Melhor os skills ali. Seria a única coisa. Que eu ia tentar. Inicialmente. Porque você vai. De cem mil funcionários. Profissionais. Para vinte. Dez. Às vezes. Dependendo da ferramenta. Então. Agora. Isso aí. Tem que fazer sentido. Para a empresa. Sim.

Entrevistadora: Tem que fazer a empresa. Parar da regra de negócio.

Francisco: Isso. Se faria sentido. Ela ia anunciar. Algumas ferramentas. Ou se ela nem entende. Do assunto. E realmente. Só precisa. De um cargo mesmo.

Entrevistadora: Tá. Aí você. Nesse caso. Você procuraria a empresa. Para bater um papo. E entender melhor. O que que era. Tá.

Francisco: Exatamente. Exatamente. Melhorar os dados. Ou seja. Tentar. Melhorar. O dado de entrada. Inicialmente. Já pela. Manualmente. Já ia querer. Tentar melhorar. O dado de entrada. Sem. Sem prosseguir. Antes de verificar.

Entrevistadora: E aí. Então. Você já. Pensando aí. Como você. Deve. Deveria. Saber mais. Para melhorar os dados. É. Você. Consegue. Imaginar. Quais métodos. Você escolheria. A partir daí. Se existiria. Algum desafio técnico. Né. Que você consegue. Já antecipar. Mensurar. E como você faria. Essa integração. Para a implementação. Né. Como foi. Implementar. Para integrar. Esse. Algoritmo. No sistema de divulgação.

Francisco: Também. Perfeito. Né. Bem. Bem. Primeiramente. Eu acho que seria. A colher. Coletar os dados. Né. Escolher. Escolher alguma plataforma. Para colher. Para coletar os dados. E. E supervisionar. Esses dados. A distribuição. Desses dados. Né. Visualizar. Gerar algumas visualizações. Desse dados. Com essa vaga. Utilizando ela. Como exemplo. Né. Porque. Assim. Quando a gente. Está falando de fazer um algoritmo. Eu não estou fazendo. Só para essa vaga. Estou fazendo algoritmo. A qualquer outra vaga. Que vai vir. Depois. Com essas mesmas variáveis. Né. Então. Eu não estou. Indo buscar

dados. Buscar. Semelhantes. Ao que eu estou procurando. Eu estou querendo. Utilizar ela. Como input. Como entrada. Para. Para modelar. Um sistema de recomendação. De profissionais. Que eu tenho em uma base. Então. Isso é uma coisa genérica. Então. Eu primeiro. Gerarei visualizações. E. Extrairia algumas estatísticas. É. Desses dados. Que eu selecionei. Na minha base de dados. Para lidar com problemas técnicos. É. De. Falta de dados. Né. Para. Ou seja. Para. Para. Verificação do dado. Da coleta. Né. Uma vez. O dado validado. Eu. Partiria. Para. Visualizar. Talvez. Todas essas variáveis. É. Fazer uma. Realmente. Um estudo de visualização. De correlação. Dessas variáveis. Manualmente. E. Quando eu digo isso. Não estou nem olhando. Quem é mais a variável. Eu só estou buscando. Correlação. Eu estou realmente. Querendo entender. Se. Se. Né. Se o dado ali. Se. Se essas features. São relevantes ou não. No meu dado. Aqui. É importante dizer que é um problema com esse meu dado. Porque. Eu. Preciso. Melhorar. Ele. Esse dado que eu coletei. Eu vou ter que. Eu vou ter que. É. Selecionar. A base de profissionais. Eu não tenho a informação. De quais profissionais. Deveriam ser selecionados. Eu só tenho.

Os skills. Então. É. É como você, começou no início, então a gente cai num problema que não é dizer se esse profissional é selecionado ou não, a gente cai num problema de buscar correlação entre os dados desse profissional e os dados da nossa vaga, né? Então, vamos lá. Realmente nunca fiz de algoritmo de recomendação, mas vamos lá.

Entrevistadora: Quando você fala assim, gerar visualização para extrair estatística, é justamente isso assim, é você conseguir transformar os dados ali em algo palpável para você fazer uma correlação entre eles.

Então, extrair estatística seria isso, a correlação.

Francisco: Isso, correlação, desvio padrão, média, olhar umas estatísticas básicas de cada uma daquelas features para ver se o dado está certo e para ver se faz sentido aquela especificação também, mas focando no algoritmo. Algoritmo, eu não quero me preocupar com os cinco anos que ele está pedindo, eu quero me preocupar só realmente como está a experiência dos profissionais hoje, então, como é hoje, por exemplo, o tempo de experiência das pessoas, qual é a média ou mediana, né? Extrair, olhar esse dado e ter uma noção.

Agora, quanto ao nosso problema aqui, eu não tive uma ideia ainda.

Entrevistadora: Quanto ao problema qual, específico, você falando de integral ou do desafio?

Francisco: É o do desafio, estou aqui pensando que é...

Entrevistadora: Não, não tem um problema, assim, isso aí é uma coisa, se você, eu acho que você já falou, na verdade, que o desafio seria você analisar de uma maneira mais aprofundar os dados, né? Entendendo, assim, a correlação,

levantar, é, quais são as features mais importantes, porque extrair os dados de quantos anos de experiência, isso aí você vai eliminar, né, por exemplo, uma possibilidade de candidatos que podem ser bons, né? Você focar nos dados, assim, muito restritos, né? Acho que, pelo menos, é isso que eu entendi que você falou.

Francisco: É, pois é, eu não tô pegando... Tá, é porque o nosso dado é uma base de currículos, né? É uma base de currículos, então, eu não tenho o dado que é, por exemplo, então, o meu dado não é devo contratar ou não esse cara, né? A gente não tem essa decisão pronta, a gente vai, o nosso algoritmo que vai gerar essa decisão, então, eu preciso, eu preciso criar alguma coisa que vá esquecendo aqueles, aqueles, as nossas especificações lá de 5 anos, não sei o que, não sei o que lá, eu posso, por exemplo, é, bem, é um ponto que talvez eu não fizesse na prática, mas acho que é o que a gente, a gente tem que, tem que focar aqui, que é pegar dados desses usuários, né? Todos os dados possíveis, observar, né? Verificar os dados, se eles estão corretos, estão completos e observar quais dados eu tenho por usuário, né? Então, por exemplo, eu vou ter esses que a gente enunciou, mas eu posso ter além desses dados, eu tenho, posso ter, por exemplo, quantos anos eles têm de experiência como Fullstack ou quantos anos ele teve, no total, né? Então, quantas empresas, quantas línguas ele fala, eu posso ter vários dados, quais as empresas que ele trabalhou, por exemplo, uma base de dados vai, vai, vai depender, né?

Eu posso ter, é, não sei, pensar aqui, bem, tudo, o nome do lugar que ele estudou, o nome do, o nome do, da cidade dele, o nome, né? Se a empresa que ele trabalhou antes era remota, eu posso ter outros dados, e aí, tendo muitos dados, além desse, que a gente tem, tá falando ali, pode fazer sentido utilizar a machine learning, para a gente reconhecer padrões entre os usuários, né? E tentar, e tentar alcançar alguma coisa, talvez só com esses dados, na prática, não fizesse sentido, um simples Excel, você consegue filtrar, tá bom, quantos tem menos de cinco anos, quantos tem mais, não sei o que, você filtra, você fica lá com seus, você não precisa de machine learning, mas pensando aqui, no nosso propósito, que é falar sobre os algoritmos de machine learning, só faz sentido, se eu, se eu tiver, quer dizer, na minha, na minha concepção, só faz sentido, se eu estender os dados para dados adicionais, dados não diretamente relacionados à vaga, então, a pergunta é, eu quero encontrar pessoas que eu tô buscando exatamente esses dados, ou eu quero encontrar pessoas que, possivelmente, são interessantes, com base nesses dados, né?

Então, eu tentaria fazer uma correlação entre, entre o desafio técnico, no caso, seria pegar todos os dados que eu tenho, e que eu acho que eu posso utilizar para isso, que faz sentido utilizar para isso, e aí pode ser, desde,

como eu falei, quantos anos de experiência ele tem em C++, quantos anos de experiência ele tem em Java, mesmo que eu nem saiba se essa é uma tecnologia importante para essa vaga, quantas vezes tem uma certa palavra no currículo dele, a palavra Java, a palavra Python, qualquer palavra, né? Eu posso utilizar esses marcadores, mesmo sem entender se eles são relevantes para mim, e tentar criar uma correlação entre usuários, né? Ou seja, eu vou aproximar as pessoas que são mais próximas, numericamente, né? Então, um cara que tem, pô, 10 anos de experiência, vai estar mais próximo de um cara, talvez, com 10 anos de experiência, um cara que é de Python, provavelmente, vai estar mais próximo de um cara de Python, um cara que trabalhou na mesma empresa, sei lá, na Rede Globo, ele vai estar mais próximo, de alguma maneira, porque naquela feature, o nome das empresas trabalhou na Globo? Ele também trabalhou, um trabalhou e ele também trabalhou, então, naquela feature ele vai estar mais próximo.

Então, eu vou aproximar as pessoas, numericamente, de acordo com as suas características, né? Com esses dados deles, né? Isso é uma maneira de buscar. E, a partir daí, a partir daí, quanto à integração aí, você me pega muito. Bem, eu tenho algo a dizer que eu... pode falar.

Entrevistadora: Pra você, quando você fala de aproximar, né? Criando essa aproximação visual das pessoas também, você depois colocaria um peso para saber o que é mais importante, para dar mais importância para aquilo que está sendo buscado na vaga?

Francisco: Inicialmente, não. Inicialmente, eu só iria... eu estou colocando as pessoas em caixinhas, dizendo, ah, essa daqui tem experiência, por exemplo, na minha concorrente. Enfim, eu vou colocar elas, eu vou aproximá-las, de acordo com os dados delas, sem qualquer juízo de valor, se isso é bom ou ruim, ou de valor para a nossa vaga aqui, para a nossa... de peso ainda.

Porque eu ainda não sei se é zero ou se é dez, ou se é um para cada feature. Mas eu tentaria, depois, fazer alguma experiência com pessoas. Então, eu poderia, por exemplo, dentro dessa grande matriz de pessoas, falar, eu vou contactar dez pessoas dessa grande matriz aqui de cem por cem, né? De repente, eu tenho ali dez mil pessoas que eu tenho que analisar, e eu vou contactar, e eu vou fazer entrevistas com essas pessoas, e eu vou tentar encontrar com qual o meu cliente se empatizou, ou com qual eu gostei da experiência, né? Eu tive a experiência de aprender.

E, pós isso, eu poderia utilizar pessoas similares àquela pessoa, né? Isso seria uma estratégia de recomendação possível para esse caso. Claro que, acho que isso, para vagas, talvez, isso é muito utilizado para produto, né? Então, se você tem um produto muito grande, né? Se você tem um sistema de, né?

De produto, sei lá, investimentos, qualquer coisa que você quer oferecer para as pessoas, você mapear pessoas similares a ela, né? Em termos de tudo, né? Em termos de dados demográficos, salário, você pode acabar fazendo uma recomendação mais certa, né? Por exemplo, se a pessoa gostou disso aqui, provavelmente ela vai gostar disso, um sistema de venda de produtos, por exemplo, né? Então, se você gostou da bola de futebol, você pode gostar da rede de futebol também, né? E como eu vou saber quem joga futebol, né? Provavelmente quem está nas cidades de praia, e aí vai. E aí você, com os outros dados, você pode alcançar mais facilmente pessoas similares, né? Para vagas, seria bem diferente, eu acho, mas como é o nosso caso aqui, estou tentando adaptar, entendeu?

Entrevistadora: Sim, mas por que você acha que seria diferente?

Francisco: Porque eu não vou oferecer até a mesma pessoa para tantos clientes.

Entrevistadora: Entendi.

Francisco: É, por isso que eu acho que a palavra recomendação aqui, a gente deveria inverter, talvez, para a vaga. Mas é porque aí depende, eu não sei exatamente como funcionaria a nossa empresa fictícia.

Vamos imaginar que, na verdade, a nossa empresa, ela é... Ela não é, o trabalho não é para a empresa que está contratando, mas para as pessoas que estão oferecendo o seu currículo. Aí pode até fazer sentido eu oferecer essas... Mas é porque uma vaga, quando ela é preenchida, ela já está preenchida, né? Não é um produto que eu vendo várias vezes, né? Então, talvez por isso, eu não... Na prática, eu não faria isso, mas com o dado que eu tenho aqui, é uma forma de visualizar, assim, uma busca, né? Uma busca pelos profissionais similares, né?

Entrevistadora: Entendi.

Francisco: Talvez não fosse prática no mundo real, mas para a gente aqui, talvez não faça essa diferença.

Não é o que a gente está querendo julgar, né? Então, ou seja, eu não sei as features, qual a importância delas. Eu não sei, por exemplo, nesse caso, a idade, se ela é relevante ou não. Para mim, é bom ser novo, mas pode ser que eu encontre pessoas... Vamos lá, por exemplo, se a minha busca é por Cobol, vai ser difícil eu encontrar, muito difícil eu encontrar alguém muito jovem.

Um linkedin, né? Eu vou achar perfis mais próximos e, naturalmente, a idade vai ser uma coisa relevante, né? Então, é bom também, antes de utilizar essa matriz de dados, você verificar se os dados são permitidos utilizar, né? Se faz sentido você colocar eles ali nessa cesta, né? E quais dados você deve ocultar, utilizar e quais dados você não precisa nem mesmo ocultar, né? Por

exemplo, identificar quem já é conectado com outra pessoa, né? Ah, meu dado diz aqui quais profissionais têm um ID nos amigos do outro. Então, esse dado eu posso utilizar, né? Então, isso tudo tem que ser... Eu posso, se eu não tiver o nome da pessoa, por exemplo, ou não tiver nenhum documento da pessoa, nada que identifique ela, né? Então, aí vai dar regulação de cada local, né?

Entrevistadora: Sim. Tá. Daí, a próxima fase, até assim, são duas coisas que se conectam, tá? Primeiro, assim, como você faria os testes e a validação, né? Para saber se tem algum tipo específico de teste que você acha que dá para validar esse algoritmo. Quais seriam as métricas de sucesso que você vai usar para avaliar o desempenho do algoritmo, até porque, né, você vai estar vendendo o algoritmo como um produto, né? Sim. E possíveis feedbacks e interações que você acha importante ter internamente, externamente, para poder... Que podem levar a alguma modificação do algoritmo, melhoria, né?

Francisco: Uhum.

Entrevistadora: Tá? E vou só antecipar a próxima, porque isso aqui meio que eu tenho visto que está se unificando. No final, assim, resultados e impactos. A efetividade tem a ver também com essa parte das métricas, né? Mas como você avalia a efetividade do algoritmo na recomendação de candidatos com os skills desejados, que acho que é um pouco, né, disso que você está bem preocupado, assim, né, que você falou, de conseguir conectar, entender melhor as skills das pessoas com as skills desejadas e o impacto na contratação, né? Como pode gerar algum impacto significativo, de fato, no processo da contratação, que isso tem alguma coisa a ver também com a parte das métricas de sucesso, tá? Vou voltar anteriormente aqui.

Francisco: Tá bem. Bem, de testes realizados, eu iria propor algumas experimentações, né? Com essa primeira versão do modelo, eu iria propor fazer algumas experimentações com novos dados, né? Com outras vagas, né? E observar se esses profissionais, em que etapa eles chegaram do processo. Eu também iria, para isso, eu ia oferecer também, eu ia sugerir que também pessoas com modelos mais simples de algoritmo, né? Fossem selecionadas, ou seja, se eu pegasse esses dados ali e eu fizesse um modelo simples, tradicional, que não fosse um algoritmo de inteligência artificial, né? Que não tivesse neurônios, eu ia propor comparação do nosso modelo de inteligência artificial, que já que é o que a gente está vendendo aqui, que precisa de inteligência artificial para fazer isso, né? Eu ia propor com modelos mais simples, como uma regressão linear, alguma coisa matemática básica de mais de 100 anos.

E aí, eu ia comparar as métricas de sucesso entre esses modelos, né? Quanto às métricas de sucesso, bem, aí eu acho que seria, é, a métrica de sucesso vai depender da nossa estrutura com o cliente, da nossa conexão com

o cliente, mas eu diria que poderia gerar algumas métricas em relação ao FIT com uma pré-seleção de currículos, mas é que a gente não tem o dado que geralmente você tem para fazer a modelagem, né? A gente está falando de um caso real, que é o que acontece nos casos reais, a gente não tem o dado que a gente quer. Então, eu teria que pensar em alguma pontagem de falsos positivos, de falsos negativos, e basicamente armazenar todas as nossas inferências. Eu focaria aqui em ter um rastreamento de todas as inferências e de todos os modelos para que a gente pudesse explicar cada uma das decisões e entender exatamente os índices de falsos positivos e de falsos negativos, mas principalmente poder explicar cada caso.

Então, eu teria lá, não, mas nesse dia aqui, eu rejeitei esse currículo aqui, por quê? Porque essa feature aqui, em específico, foi muito negativa. Então, ter armazenado esse rastreamento. Agora...

Entrevistadora: Próximo, você quer?

Francisco: De feedback e interações? É, eu acho que aí é exatamente você experimentar o seu modelo e tentar coletar dados também de aceite, não é? Não sei se daria para a gente pensar em alguma coisa antes disso, mas... Bem, eu acho que aí cada um tem a sua interpretação.

Eu teria de fazer uma pesquisa com o entrevistado para entender se ele achou que aquilo era relevante para ele, aquela posição, enquanto também a quem fez as etapas, assim. Ou seja, criar um pipeline aí de ter um... Assim como na métrica de sucesso, eu ter também registro de todas as minhas previsões, ou seja, todas as decisões que foram tomadas. Eu tenho um registro de tudo o que aconteceu com ela no mundo matemático, mas também, ou seja, dos pesos ali, mas também ter o registro do que aconteceu após ela, em efeito dela, para tentar chegar a alguma conclusão.

Por exemplo, dizer, olha, nós tivemos aqui... 10

Entrevistadora: E seria um feedback mais humano mesmo, não é? Assim, uma questão de você pegar com pessoas já a impressão da utilização, do que você...

Francisco: Depende. Nesse caso, sim, mas se eu estivesse aplicando isso para um produto, por exemplo, eu estou selecionando as pessoas que eu vou recomendar, grandão, assim, compre esse produto porque os últimos parecidos com você compraram.

Se ela falar não, obrigado, dali eu já posso coletar essa informação, não é? Mas no nosso caso aqui, como é uma questão humana, eu acho que é... Provavelmente, sim, seria coletar esses dados de maneira humana, ou mandar um formulário, uma pesquisa, alguma coisa assim.

Entrevistadora: Tá. Tá bom, aqui é aquilo que eu falei, né? Assim, o

impacto na contratação, acho que você já falou também, né? Em relação até à questão da métrica de sucesso, não é? Mas se quiser falar mais alguma coisa.

Francisco: É, eu não tenho nenhum insight maneiro sobre isso. Eu pensaria que, realmente, é talvez... Acho que a única coisa a somar nisso que eu tenho é sempre comparar ele com alguma análise mais simples, né? Já que a gente está fazendo um algoritmo que é artificial, que a gente não tem controle

Então, quando alguém perguntar assim, por que eu uso um algoritmo artificial? Falo, porque se a gente fizesse uma regressão linear simples, estaria tendo esse percentual de acerto aqui, a gente está tendo este aqui. Então, acompanhar isso, ter observabilidade, eu diria. Eu acho que, como eu vou fornecer esse serviço, eu ia me preocupar em trazer observabilidade sobre a efetividade desse serviço no tempo, na prática, né? E para que isso seja uma métrica também do meu produto.

Eu dizer, olha, eu aumento em tanto, eu alcanço uma efetividade de tanto. Quanto ao impacto na contratação, a expectativa é o que gera o impacto na contratação, né? De qualquer maneira, quanto ao impacto, eu acho que é sempre importante que seja, também tem observabilidade sobre o impacto, sobre o que está sendo decisivo, trazer insights, o que está sendo decisivo para aquele modelo. Então, quando eu digo que, pô, eu estou se recomendando essa pessoa, eu não vou dizer só isso, eu vou dizer, eu estou recomendando essa pessoa, porque nesse critério aqui, ela combinou muito com o que eu questionei para o modelo, né? Ou seja, sempre ter alguma informação extra.

Então, eu diria que é isso, assim, para que você possa avaliar impactos positivos ou negativos, não sei, muitas vezes isso pode explicar e trazer informação, mas eu não tenho muita experiência nisso.

Entrevistadora: E o impacto na sociedade como um todo, né? Assim, no reflexo da contratação, pensando nesse âmbito mais geral.

Francisco: Bem, eu diria que, da mesma maneira, mesma maneira não, eu diria que um algoritmo, ele requer um cuidado extra, justamente por ser uma caixa preta, ele requer um cuidado extra na sua observação, na observabilidade, porque muitas vezes, o mesmo algoritmo pode não mudar, mas a sociedade pode mudar.

Então, você tem lá uma coisa que funciona há 10 anos daquele jeito, e ela se comporta de uma maneira útil para a sociedade. A depender da mudança da sociedade, aquele mesmo comportamento, se você reproduzir ele em uma outra sociedade, num outro momento, ele pode ter um comportamento nocivo para a sociedade. Então, eu diria que a observação, ela é contínua, ela tem que ser contínua nesse sentido, né? Uma tendência nos dados pode fazer um

algoritmo igual, estático, a se tornar tendencioso, que ele não era, mas ele pode começar a ser tendencioso, a depender de mudanças que podem acontecer com políticas públicas, com tendências de tecnologia, mudanças na sociedade, e que, de repente, podem se comportar de um jeito diferente.

Então, eu diria isso, eu acho que o impacto, eu acho que é o cuidado de aplicar o algoritmo, entendendo sempre para que ele está sendo utilizado. Por exemplo, falsos positivos podem gerar problemas, então, de repente, eu tenho que focar mais nos negativos, então, eu acho que é sempre saber, você não faz um algoritmo sem entender o impacto dele.

Entrevistadora: O contexto, né?

Francisco: Na empresa, no sistema, na sociedade, o que ele está fazendo? Ele está detectando uma pessoa, né? Para quê? Para identificar na entrada da casa dela, pô, a entrada da casa dela é uma coisa muito crítica, então, você pode ter falsos positivos na entrada da sua casa, você não quer falsos positivos, você prefere que você não seja reconhecido do que reconhecer qualquer pessoa e dizer que sou eu.

Então, eu prefiro, de repente, ficar mais tempo ali, coletar mais dados para garantir que sou eu.

Entrevistadora: Entendi.

Francisco: Então, eu acho que o algoritmo não nasce nos dados, ele nasce na aplicação, justamente, com ele.

E eu acho que é aí que, geralmente, pode haver um... podem haver exageros ou também displicências quanto ao impacto, né?

Entrevistadora: Uhum, legal. Quando você fala observabilidade, isso é um termo que está sendo usado, um termo técnico?

Francisco: Sim. É ter... Como é que é isso? É ter conhecimento observability, que é, exatamente, você ter visão sobre o seu algoritmo, sobre o seu sistema, né? Vou dar um exemplo para você, para a gente aqui.

Se eu tenho um sistema que me manda um SMS, quando, sei lá, a câmera da minha casa pega um rosto que não é meu, e eu moro sozinho, eu posso ficar na minha casa por um ano sem receber visita. E esse sistema não vai falar nada comigo durante um ano. E aí, quando alguém entrar, eu vou esperar que ele funcione? Então, ter observabilidade é você ter uma maneira de observar se tem funcionado.

Ele tem me visto, ele tem, né? Ele tem falado para mim, semanalmente, que essa semana você não recebeu nenhuma visita, por exemplo, né? Ou seja, ele está em contato com o usuário ou com o administrador e trazer uma segurança de que ele está a funcionar naquele período. Então, seria mais isso. Ou seja, não abandonar uma coisa funcionando e esperar que ela funcione.

Então, acho que o modelo, qualquer modelo de inteligência artificial, ele merece, ele precisa ser observado, assim. Principalmente, vai depender, principalmente, do tipo de impacto que ele gera, né?

Entrevistadora: Tá. Muito bom, Francisco. Eu acho que foi isso. Não sei se você quer colocar mais alguma coisa, mas aqui a gente acaba.

Francisco: Não, eu não tenho muito a somar, não. É só isso mesmo.

Entrevistadora: Tem sim, já falou um monte de coisa, ué. Tá bom. Deixa eu ver aqui como é que eu paro de gravar, rapidinho.

B.6

João

Entrevistadora: Senão, fica uma coisa assim... eles ficam tentando colocar... deixa eu ver se tá certo... tá certo já. Então, no outro dia eu fiz uma entrevista em português e a transcrição estava em inglês, ficou uma coisa maravilhosa.

João: Eu nem sabia que dava para fazer transcrição pelo Teams, que ótimo, né?

Entrevistadora: Tá, e não é ruim não, tá? É boa, é boa.

Tem outras ferramentas mais, assim, assertivas ali, né? Mas essa é boa também. Deixa eu te falar como é que vai ser o esquema e o nosso objetivo aqui também. Eu vou compartilhar com você um PowerPoint que eu fiz só para gente... com roteiro, tá? A ideia é a gente fazer uma conversa baseada num método de design, chama Directed Storytelling, que é... eu vou pedir para você pensar, desenvolver o raciocínio em cima de um contexto, né? De uma atividade.

E aí você vai me falando o que você faria. Se você quiser, por exemplo, ah, vou pegar... eu pegaria a ferramenta tal. Se eu quiser me mostrar essa ferramenta, fica à vontade, não precisa, tá? Mas não é obrigatório, quer dizer, mas fica à vontade.

Mas vou compartilhar aqui para gente, assim, funcionou usando esse roteiro como fundo, porque acho que as pessoas geralmente são visuais, né? Então é mais fácil para você ir olhando e não se perder. Mas se você quiser, a hora que você quiser, eu tiro aqui, você compartilha sua tela e a gente vai desenvolvendo essa conversa. Primeiro eu queria saber um pouco mais, né? Eu já fiz algumas perguntas lá no formulário, mas queria saber um pouco da sua experiência com algoritmos, com inteligência artificial, algumas questões de experiência mesmo que você já teve de trabalho, ou então na faculdade, mestrado, você pode trazer aqui.

Queria entender um pouco melhor isso.

João: Então acho que vale trazer essa contextualização contando o meu caminho, né? Até chegar aqui, acho que eu fiz faculdade de engenharia elétrica, mas logo no começo da faculdade, no segundo, terceiro período, eu comecei a tocar um projeto de pesquisa, iniciação científica, dentro do Instituto de Computação. Isso acontecia dentro do Instituto, do Laboratório de Computação Gráfica.

Então, foi nesse momento que eu comecei a codar já escrevia códigos, mas ter um conhecimento maior sobre computação gráfica e a partir disso escrever algoritmos. E aí eu vou caminhando na universidade e o primeiro paper que eu escrevo, ainda estou na graduação, é sobre o desenvolvimento de junção de dois algoritmos para fazer a reconstrução de métodos de imagem 3D. Então é uma coisa bem técnica, são algoritmos, mas não são algoritmos de machine learning.

São de computação gráfica para poder fazer processamento de imagem. Eu começo a estagiar, acho que em 2017, 2018, e aí eu começo a estagiar numa empresa que é a Radix, e lá eu começo a trabalhar diretamente com machine learning, aplicando soluções para a indústria. Eles também são uma empresa que presta consultoria e projetos, então eu comecei já a aplicar técnicas de machine learning, e aí técnicas clássicas e o que tinha de mais avançado também na época, para resolver problemas de clientes, soluções industriais.

Depois disso, eu troco para a Deloitte, eu estagiava numa empresa de elétrica, mas nada na área, comecei a estagiar na Radix, depois eu fui efetivado, trabalho lá mais uns anos, fazendo o mesmo desenvolvimento, e aí passo por diversos tipos de desenvolvimento de algoritmos. Processamento de imagem, análise de dados, chatbot, coisas do tipo, e aí eu vou para a Deloitte, e continuo esse caminho. Hoje, meu cargo é engenheiro de visão computacional, consultor, mas eu trabalho, acho que, seguindo um pouco a mesma lógica que acontecia quando eu era da Radix, onde a necessidade do desenvolvedor ou de uma pessoa que escreva algoritmos e que trabalhe com machine learning de uma forma geral, eu sou colocado.

Então, isso envolve visão computacional, mas também envolve GMA e para chatbots e qualquer área que seja relacionada a desenvolvimento de algoritmos e machine learning.

Entrevistadora: Legal. Essa parte de algoritmo de recomendação, especificamente, você já chegou a trabalhar, construir alguma coisa?

João: De recomendação?

Entrevistadora: Aham.

João: Não, acho que o mais próximo que eu cheguei foi chatbots.

Entrevistadora: Tá. O que eu vou propor hoje é a gente fazer uma

simulação que tem a ver com esse universo do algoritmo de recomendação.

Mas, assim, não tem problema você não ter desenvolvido antes. A minha intenção aqui é investigar o seu racional mesmo, como você construiria essa solução, afinal, é uma solução. Então, o objetivo, o nosso objetivo é o seguinte.

Num cenário fictício, em que você trabalha numa empresa que faz algoritmos, modelos, uma empresa de tecnologia de porte-médio te contrata, contrata a sua empresa para fazer uma divulgação de vagas, para anunciar uma vaga em tecnologia. Então, é uma empresa A, empresa B. A empresa A, de tecnologia, contrata uma empresa B, que vai construir um anúncio para conseguir pescar possíveis talentos com algumas skills necessárias para aquela vaga. E aí, qual é a premissa? Quais são as definições que essa empresa de tecnologia passa? Algumas, poucas, que a vaga é de tecnologia para um profissional pleno, que ele deve ser um desenvolvedor full-stack ou desenvolvedora, com uma formação em tecnologia, engenharia de software, sistemas ou algumas outras formações afins, pode, né, outro tipo de engenharia também, mas da área de STEAM e a experiência de cinco anos de experiência, pelo menos.

No regime remoto, tá? Então, essa, a empresa chegou e contratou e passou isso para vocês, não passou mais nenhum detalhamento, tá? Então, a ideia é eu pensar, você queria que você me contasse, né, o como você faria. Primeiro, no âmbito do planejamento e a concepção dessa solução, né, assim, que seriam os requisitos, como você definiria, como você buscaria a definição desses requisitos e quais seriam os dados utilizados, né, ou que tipo de dado você precisa, como que você pretende coletar e processar esses dados?

João: Parece um escopo bem aberto, né? Então, para fazer a definição dos requisitos, eu acho que os requisitos são só os requisitos que foram colocados na vaga, né? Então, como eu definiria eles? Acho que a partir do que foi dado como informação. Tem informações gerais, então a busca vai ser mais ampla, né? E os dados? Que tipo de dados você utilizaria para alimentar esse algoritmo? Como esses dados foram coletados e processados? Eu não sei muito bem se eu consegui entender o objetivo, não o objetivo, mas como assim? Você quer uma resposta técnica sobre como eu faria essa análise?

Entrevistadora: Eu vou citar um exemplo, mas não é para você se limitar, mas, por exemplo, até voltei aqui nessa tela de objetivos falando do cenário, né? Como que você acha? E aí, claro, a gente está num cenário aberto, numa simulação, vamos dizer que você procuraria no LinkedIn ou colocaria, montaria esse anúncio para o LinkedIn ou Facebook, qualquer uma dessas plataformas.

E você, como que você atrai pessoas para uma vaga de tecnologia com o perfil de profissional pleno? Quem seriam essas pessoas? Como seria, assim,

você, como você define os dados que são de interesse para você saber, né? De como é que você entende, né? Busca aí dentro de perfil profissional de uma pessoa, que ela é desenvolvedora ou não. Entendeu, assim, mais? Sim.

João: É, eu acho que eu buscaria entender as qualidades técnicas dela, né? As capacidades técnicas.

Então, acho que um fator a ser considerado a partir desse objetivo dessa vaga é o histórico, e aí pode ser incluído o histórico profissional de experiência e o histórico de currículo escolar, assim. Acho que a partir disso você consegue informação pertinente. Eu não consigo enxergar outra ferramenta de divulgação para a procura dessa vaga para além do LinkedIn.

Eu acho que, quer dizer, consigo enxergar, mas acho que essa pode ser a mais interessante, assim, porque é onde existe o mercado profissional hoje, assim. Esse não é especificamente minha área, né? Não saberia como fazer isso da melhor forma, mas sim, acho que o processo de histórico contaria, histórico profissional, e aí isso pode ser destrinchado, né? E aí, como o alcance pode ser feito a partir disso? Quando é pensado um tipo de algoritmo de recomendação ou algo do tipo, eu sempre penso numa estrutura de gráficos, né? Que você consegue fazer uma análise de conexão a partir de conexões pré-existentes e do que você tem em comum a partir disso, né? Então, seria uma análise parecida, pensando numa estrutura de gráficos e colocando como, não sei, talvez as skills que existem em relação à vaga e esses pontos de conexão entre outras pessoas também. Seria algo mais ou menos assim.

Entrevistadora: Meio que para ir vendo, assim, o relacionamento dessa pessoa com outras pessoas da área. Seria isso?

João: Isso influencia, acredito que isso influenciaria.

Eu não sei se isso seria um tomador de decisão, mas sim, acho que no processo de divulgação, pensando na lógica de divulgar a vaga.

Entrevistadora: Tá. Aí, assim, não precisa ficar focado na parte técnica em si, como faz a conexão com a plataforma, mas é mais assim, o que você daria um peso, na hora de dar um peso para procurar essas características, o que seriam características dessas pessoas que você considera que seriam importantes? Então, você já falou, histórico profissional, né? Histórico...

João: É, mas eu não sei se histórico é uma característica pessoal, né? Eu acho que, sei lá, características pessoais que eu colocaria são as interpessoais, principalmente, né? Capacidade de relacionamento, como essa pessoa se adequa ao ambiente corporativo, considerando que seja uma vaga de perfil corporativo.

E essa experiência profissional influencia por conta do perfil da vaga, né? Acho que pede uma pessoa sênior e acho que isso sempre envolve experiência, né?

Entrevistadora: Tá. Então, deixa eu passar para... Você conseguiria pensar na escolha de algum método, né? Algum algoritmo específico, sei lá, árvore, outro algoritmo, né? Outro método, para fazer a definição, usar essas features que você vai fazer, né? Essa definição para rodar o algoritmo. Você consegue pensar uma parte... Não precisa pensar na integração, implementação, como falei, né? Tá muito aberto essa questão de como seria a plataforma, né? E se teria algum desafio técnico no retorno desse... pensando nos resultados mesmo.

João: Entendi. Acho que a primeira coisa é classificação, para considerar o que entraria ou não entraria dentro dos requisitos mínimos. Porque se você tem candidatos que não atendem os requisitos mínimos colocados, que são mais amplos, mas que não atendem eles, acho que não entraria.

E aí, a partir disso, talvez, clusterização. Conseguir clusterizar esses grupos, essa população de indivíduos a partir dos principais skills deles. E aí conseguir trazer um panorama, um mapa mais definido para o cliente, né? Para quem tem interesse em fazer essa contratação.

Por exemplo, você pode ter profissionais que têm uma ênfase em uma determinada área que é pertinente a esse escopo aberto da vaga, e outros que têm uma ênfase maior em outra área que também se enquadraria nesse perfil mais amplo. Então, acho que fazer um processo de clusterização para entender onde está essa população de indivíduos faria sentido para mim.

Entrevistadora: Tá. E você teria, assim, agora até a nossa outra parte aqui, em relação aos testes e validação. Você vai rodar, né? E vai entender, assim, o que realmente é um sucesso, né? Qual foi o resultado que foi retornado? E em relação, assim, a gente está falando de pessoas que vão, que vocês vão atrair com esse algoritmo, né? Assim, que vai chegar esse anúncio. Versus, né? Essa escolha, essa seleção versus o quanto realmente a empresa de contratação vai aproveitar, né? A empresa de tecnologia vai aproveitar.

Então, assim, foram selecionados 200 candidatos e não foi nenhum contratado, né? Não sei como você poderia fazer essa medição de entendimento do que você está selecionando está sendo positivo, está sendo assertivo.

João: Como fazer uma métrica aí, supondo, né? Para entender se essa análise e classificação, essa análise de candidatos está sendo efetiva para o cliente, se é o que ele quer. Essa é uma boa pergunta, porque considerando uma análise de classificação, a gente não chegaria com o resultado definido, mas sim com o mapeamento do que foi feito antes, né? Mapeamento dos candidatos.

Quais tipos de teste seriam realizados para validar o algoritmo? Teste A-B? Não sei se teste A-B é uma boa para fazer essa validação. É, eu acho que dentro desse escopo não seria aplicado nenhum teste de validação. Seria muito mais entendimento do que faria sentido para o cliente, mas acho que

simplesmente não teria nenhum teste de validação nesse momento, pensando nessa lógica que eu estou trazendo.

Quais métricas seriam utilizadas para avaliar o desempenho do algoritmo? É, acho que nesse sentido o que avalia desempenho é se a empresa vai contratar ou não algum desse cliente, né? Então também não teria uma métrica de porcentagem, um valor, sei lá, um valor não binário para poder fazer essa métrica de sucesso. Acho que seria mais se houve ou não a contratação a partir dessa população e desses indivíduos mostrados. Como avaliar feedbacks interno e externo que podem levar a modificação do algoritmo? Tá, como avaliar o feedback?

Entrevistadora: Posso te fazer uma colocação? Pode, de repente, pode te ajudar? É, como eu coloquei lá a definição do cenário, né? Lá no começo, a empresa que contratou passou, sei lá, quatro requisitos assim bem abrangentes, gerais e tal.

E aí você começou, você é um cara, né, que não trabalha, você não trabalha no RH, você não é um especialista em contratação, você vai fazer um modelo que vai rodar, enfim, com esses requisitos. Você acha, né, assim, ou diante de outro que já passou, você acha que o envolvimento de pessoas do negócio para fazer a definição das regras de negócio seria importante? E aí nessa parte do feedback, será que teria como fazer, envolver, você consideraria envolver pessoas do negócio específico, pessoas de RH, pessoas da empresa, enfim, que entendam do negócio para fazer essa definição, né, e entender melhor, trazer melhor um feedback para você e de alguma forma você ajustar o modelo? O que você acha?

João: Quando eu penso em feedback e nesse processo de interação, eu penso muito mais numa análise técnica para conseguir fazer um mapeamento desse conjunto de dados que eu tenho, dessa população, do que especificamente pensar numa área de recrutamento de recursos humanos que tem mais informações sobre, sei lá, pode ter um tato maior sobre um processo de contratação, específico. Então, acho que esse feedback para mim envolveria muito mais uma análise técnica de como seria feito essa possível clusterização, essa possível análise de grafos.

Então, seria muito mais com pares do que outra coisa.

Entrevistadora: Entendi. E essa questão, quando você falou assim, fazer análise dessa população que está retornando, né, assim, você se preocuparia, né, você pegou os requisitos, fez a definição do algoritmo e retornou a uma determinada população ali.

Em cima disso, fazendo um exercício assim de realmente imaginar, você analisaria algumas questões principais dentro dessas pessoas que estão

retornando? Tipo, características da pessoa ou para ver até se está sendo justo, né, se essa seleção está sendo justa?

João: Eu acho que tem algumas questões aí nessa pergunta. Considerando, acho que de uma forma geral, a estrutura social que é criada e que existe, a partir disso, como se dá as características e as pessoas que ocupam as vagas, é possível pensar formas de diminuir enviesamento, né? Mas essa análise que eu estou considerando já é pensando nisso, acho que não é para ter vieses. E a partir disso, eu acho que inevitavelmente a análise da população vai sim possuir características próprias, porque se eu tenho um escopo que considera apenas os requisitos técnicos, que é o que está sendo apresentado inicialmente, num aspecto bem amplo, essa população de indivíduos provavelmente vai ter uma característica própria.

Não por conta da análise desses dados, mas por conta de como a sociedade é representada nessa população. Então, para ser tomada alguma análise mais minuciosa a partir disso, é preciso que exista ativamente a intenção de que seja feita uma análise em cima dessa característica da população. O que eu quero dizer é que é preciso que ativamente a empresa ou o contratante entenda esses... não são vieses, mas entendam que essa população vai estar representada a partir do que a nossa sociedade representa.

Então, provavelmente não vai existir nenhuma igualdade de gênero, nenhuma igualdade racial dentro dessa população. Mas aí, se a pessoa, se o contratante quer que isso seja considerado de alguma forma, que seja minimizado, ou seja, sei lá, tem alguma característica mais específica, aí pode ser considerado, por exemplo, sabe? Mas, pensando no cenário que é me colocado, é possível perceber que vai haver, na amostragem dessa população, ela não vai estar representada como é a população brasileira, ou de qualquer lugar do mundo, mas está sendo feita já a partir de um recorte, né?

Entrevistadora: Sim. Tá. Por fim, é um pouco disso, em relação ao que a gente já falou, né? Essa parte da efetividade, né? Dessa relação entre a recomendação e a efetividade da contratação, e o impacto na contratação, de fato, né? E eu acho que você já falou um pouco sobre isso, sobre ter um reflexo de como é a organização da sociedade, mas é um negócio, né? É uma solução que foi contratada por uma empresa, depende também do direcionamento deles, para você, como pessoa técnica, fazer esse direcionamento, né? Foi isso. Não sei se você quer colocar mais alguma coisa nesse contexto, alguma observação?

João: Ah, eu acho que, ativamente, a gente tem que ter essa percepção e fazer, sabe? Sempre uma análise que considere a sub-representação das pessoas em diferentes espaços da sociedade, ou esse recorte. Então, sei lá,

pensando nesse cenário que a gente trouxe aqui, eu acho que o imparcial não necessariamente é o melhor caminho, né? Eu acho que isso é um ponto também, né? É importante a gente ser crítico, e não só imparcial.

Pensando num cenário técnico, completamente técnico, é possível, acho que os algoritmos, de alguma forma, têm que ser imparciais, mas a nossa leitura e o nosso posicionamento contra a pessoa, não, né? E aí, a partir desse lugar técnico que eu exerço, muitas das vezes, é preciso que o meu entendimento de mundo faça com que eu crie algoritmos imparciais. Acho que é isso.

Entrevistadora: Muito bom, João.

Muito, muito obrigada, viu? Foi ótimo. Já vai para eu tirar bastante impressões aqui, inputs para minha pesquisa. E, mais para frente, quando eu começar a ter resultado, depois eu posso ir compartilhando também com vocês.

João: Ah, sim. Obrigado. E qualquer coisa, qualquer dúvida, qualquer questionamento, estou aberto também para a gente poder conversar mais.

Entrevistadora: Muito obrigada mesmo.

João: Eu que agradeço o convite. Boa continuação de trabalho.

Entrevistadora: Obrigada. Tchau, tchau.

B.7

Pedro

Entrevistadora: Vai direto para o inglês. Deixa eu te falar, qual que é a proposta? Eu vou passar uma, é uma entrevista semi-estruturada, mas a gente também chama, a gente chama de ir passando, né, por seu mapa mental, em relação a uma solução que você vai precisar construir, tá? Então, é muito importante que você vá falando, pensando, mesmo que seja mesmo confuso no começo, e se você quiser mostrar alguma coisa, alguma ferramenta, alguma plataforma, alguma coisa que você acha que dá para estruturar melhor, o que você está dizendo, fica à vontade. Eu vou compartilhar aqui a minha tela, só porque, assim, fica mais fácil quando a gente vê, acho que a gente assimila mais, né, assim, exatamente o que a gente tá falando.

Então, eu vou pedir, a princípio, para você fazer uma contextualização, até para eu entender melhor, assim, qual a sua experiência em relação, se você já teve alguma experiência com algoritmos, e qual é, especificamente, a sua experiência em relação a isso?

Pedro: Tá, vamos lá. Aqui na Deloitte, eu tô desde 2021, tá? Ficou contratado para atuar como cientista de dados. Então, assim, desde lá para cá, eu trabalho com modelagem, usando modelos machine learning, algoritmo de modo geral, tá? E aí vem algoritmo para textos, classificação de textos, reconhecimentos de identidades, também em textos, tá? Então, reconhecer

algumas categorias ali, escritas no texto, e mais recentemente, vem agora esse boom de agenda narrativa, né? Então, eu tô trabalhando em dois clientes, muito pesado nisso, tá? Mas já trabalhei com modelos para séries temporais, modelos mais para dados tabulares ali mesmo, então, ah, quero classificar nível de falha ou não falha, alguns equipamentos, coisas desse tipo, tá?

Entrevistadora: Tá ótimo.

Então, assim, você nunca teve experiência com algoritmos de recomendação, alguma coisa assim?

Pedro: Recomendação, não, nunca cheguei. Ou na Deloitte, ou fora da Deloitte, nunca trabalhei com isso, tá? Não cheguei nesse momento. Tá.

Entrevistadora: Vou te passar o cenário aqui, e aí eu queria que você me dissesse qual o seu caminho, né? Um caminho que você seguiria. Imagina que uma empresa de tecnologia de porte médio contrata outra empresa para fazer divulgação de uma vaga, tá? E você é da empresa que vai ser contratada. Então, um cliente, né, que é uma empresa de tecnologia de porte médio, te contrata para anunciar a seguinte oportunidade.

Passa só esses requisitos, né? É uma vaga de tecnologia para um internacional pleno. Skills, colocou desenvolvedora ou desenvolvedora full stack. Formação, engenharia de software, tecnologia da informação, sistemas e afins, né, área de tecnologia.

Experiência mais ou menos cinco anos e regime remoto. Tá, passou só isso, não passou mais informações. Então, a primeira etapa é como você faria o planejamento e a concepção dessa solução, né? Assim, em relação à definição dos requisitos, como é que você melhora ali a definição dos requisitos, né? Como é que você consegue fazer estruturado isso e quais os dados você utilizaria.

Pedro: Tá, vamos lá. Eu mandaria algumas perguntas para esclarecer um pouco mais sobre essa vaga, porque de fato, apenas isso é muito abrangente. Tem limitação de... por mais que seja remoto, existe limitação de alguma localidade que impeça a contratação? Ou então existe um foco maior em algum local? Seria, acho que, a primeira pergunta.

Pessoas que estão estudando, por exemplo, tem algum grupo de pessoas ali e está terminando a faculdade, mas tem interesse contratar esse grupo? Ou só pessoas que já estão formadas? É obrigado ter formação, entendeu? Então, eu iria mais nesse ponto. A questão de experiência, tá? Cinco anos de experiência, mas em alguma tecnologia específica, de modo geral? Ia esclarecer mais esses detalhes, tá? Então, eu ia muito mais por isso. E, como fonte de dados, LinkedIn, acho que é o principal, eu iria também consultar Kaggle, e tem uma comunidade de dados, mas acaba envolvendo todas as áreas de tecnologia, que se chama Data Hackers.

Eles divulgam um relatório, muito focado na área de ciência de dados, engenharia de dados, essas coisas, mas eles divulgam também algumas habilidades para a Full-stack e tudo mais. Então, ali dentro, é um Slack, um canal de Slack, aberto, assim, qualquer pessoa pode entrar. Eles falam algumas tecnologias para a Full-stack, que algumas vagas pedem.

Então, eu ia analisar, assim, outras vagas semelhantes a essa, pedem tais tecnologias, tais conhecimentos. Eu ia pegar, ali, uma coisa que está similar, muito parecida, para tentar atacar inicialmente esse problema, entendeu?

Entrevistadora: Então, você veria, assim, o que estão pedindo no mercado, já, para uma vaga similar, para trazer isso para você, tá? E aí, específicos, imagina que você está falando dos requisitos, né? Definiu melhor, assim, os requisitos, e aí, tudo bem, assim, não importa muito a plataforma real, mas vamos dizer que você vai pegar do LinkedIn. Você conseguiria dizer, assim, qual dado que você puxaria do candidato, do possível candidato, especificamente, que te interessaria para esses, para cobrir esses requisitos, né?

Pedro: Eu tentaria evitar, assim, puxar foto, nome, para identificar o gênero, por exemplo.

Eu acho que seria interessante que a avaliação da pessoa candidata, ali, fosse o mais genérica possível, em relação a gênero, etnia, religião, coisas mais pessoais, assim. Então, porque, independente do gênero, a gente tem que avaliar a capacidade da pessoa. Mas, algumas empresas, não necessariamente, olham para isso, tem muito preconceito.

Então, assim, se eu tivesse essa liberdade, eu ia dizer, ó, não vou pegar a foto, não vou pegar nome, nada disso. Mas vou pegar que essa pessoa, ela tem essas skills. Quer entrevistar com ela? Beleza.

Ela é uma mulher, um homem, e aí seguia. Mas, de modo geral, imagem, nome, experiência profissional, pegar um resumo das experiências profissionais, lá no LinkedIn, por exemplo, e agora, acho que, não sei se é recente, mas eu vi recentemente, você consegue colocar skills em cada experiência que ajudou ali. Então, pegar essas informações.

Eu gosto muito, também, de olhar quando a gente recebe aqui pessoas candidatas a vagas. Se a pessoa, a pessoa fez mestrado, doutorado, mas ela participou ativamente em pesquisa, então tem alguma publicação, eu gosto de olhar isso. Então, talvez, aqui, talvez não tanto, mas acho que pode ser um plus, uma pessoa com isso também, entendeu? Então, acho que é isso.

E, pensando mais um pouco, como é algo, esse cenário, talvez tem algum feat cultural, digamos assim, e você pode seguir algumas instituições ali no LinkedIn. Então, talvez, pegar instituições que estejam aderentes ao feat cultural da empresa, um filtro inicial, digamos assim, mas que não é nada

muito certo, né? Mas, talvez, para dizer, pô, essa pessoa tem um perfil mais para seguir esse tipo de empresas, ou tipo de ações, que minha empresa aqui vê diferente. Não sei, tá? Mas, estou pensando mais abrangente.

Entrevistadora: Para saber, para sair um pouco do classificado, assim, como mais técnico, né? Para entender mais o...

Pedro: Porque, assim, sei lá, se eu pensar uma empresa de alimentos, mas que ela defende fortemente... Não, eu estou perguntando. Uma empresa de cosméticos que defende fortemente que não deve fazer testes de animais. Por exemplo, tudo bem, aí seja essa empresa que me contratou, mas o candidato, ele segue páginas no LinkedIn que estão nem aí para isso, e se pode fazer teste animal, prejudicar animal, tá nem aí.

Tipo, pô, vai ter um feat completamente diferente para a empresa. Então, talvez, a empresa não queira algo do tipo.

Entrevistadora: Entendi. E eu não entendi quando você falou, assim, foto, nome, gênero, você não pegaria ou você pegaria?

Pedro: Eu pegaria, mas, no princípio, eu não enviaria para a empresa. A não ser que ela pedisse, óbvio, né? Mas, se eu pudesse não enviar, eu não enviaria. Mas, como, em geral, as empresas depois pedem, então, eu pegaria para ter isso inicialmente.

Enviaria apenas a capacidade, o conhecimento técnico da pessoa, depois eu enviaria essa informação, porque, em algum momento, eu ia ter que enviar. Então, para não ter que buscar de novo, eu já coletava isso inicialmente.

Entrevistadora: Entendi. Legal. Tá bom. Nessa parte, então, vamos dizer, você já definiu essa parte inicial, e aí, como é que você faria o desenvolvimento desse algoritmo, né? Assim, você tem uma ideia de qual método você vai usar? Se tem algum desafio técnico que você acha que já pode prever, e aí isso influencia na escolha do método?

Pedro: Então, vamos lá.

Como a gente quer procurar vaga, eu, inicialmente, ou iria procurar alguma base rotulada de dados, para dizer, esse candidato se encaixou bem com a vaga de full-stack, como é muito abrangente, então poderia ser algo desse tipo. Eu ia dar uma pesquisada nisso para ver se tinha. Se tiver, eu iria seguir para o lado de algoritmos supervisionados, de classificação mesmo.

Então, para cada candidato, eu classificasse se aquele candidato ou a candidata está mais aderente ou não àquela vaga. Digamos assim, eu iria por esse caminho. Qual algoritmo usar? Assim, é muito abrangente.

Eu iria iniciar pelos mais clássicos, uma regressão logística, um SVM, um Arrando Forest, são algoritmos mais clássicos, mais cotidianos, e um XGBoost. E se eles não se molhassem certo, aí eu iria para modelos mais robustos,

Deep Learning, por exemplo, ou então eu iria tentar algum algoritmo de recomendação mesmo. Digamos assim, então eu ia tentar modelar o dado para alguma recomendação.

Mas qual algoritmo de recomendação? Como eu nunca trabalhei, eu não saberia agora, eu teria que dar uma pesquisada. Mas eu iria muito para essa área. Caso não exista essa base rotulada, então a gente vai ter que partir para um algoritmo não supervisionado.

Então, um algoritmo de agrupamento de pessoas e dentro daquele grupo analisar se aquele grupo é aderente ou não. Então, seria um trabalho muito mais manual, mas também teria essa possibilidade indo para o uso de algoritmos de machine learning. Tá, entendi.

Entrevistadora: E aí, você teria como pensar também quais testes você poderia fazer para ver se está indo no caminho certo e desenvolver alguma métrica de sucesso e pegar algum feedback para as iterações, humanos, não humanos. E a métrica de sucesso até voltada para, pensando assim, você foi contratado para entregar um produto, para dizer que aquilo está funcionando.

Pedro: Tá, vamos lá, testes.

Aí, novamente, se for supervisionado, eu não usaria a acurácia, que é a métrica padrão, porque a gente vai ter um algoritmo completamente desbalanceado à base de dados da gente, porque vai ter muitas pessoas, mas pouquíssimas que são aderentes à vaga. Porque a quantidade de candidatos possíveis é muito grande, mas aqueles que realmente interessam são poucos. Então, eu iria para alguma métrica mais de precisão, recall e F1-score, algo desse tipo.

Eu ia para essas métricas, então essas métricas tendo uma performance. Eu coloco acima de 75, 80, muito por experiência de outros projetos, tá? Mas pode ser que o cliente queira algo muito acima disso, 90

Em geral, acima de 80

Se o candidato ou a empresa tivesse uma lista de candidatos, eu ia segurar essa lista para ser o conjunto de teste final, por exemplo, e pedir para eles revidarem, entendeu? Então, assim, eu ia pegar o conjunto que eles me fizeram, passava pelo modelo. O modelo iria fazer a classificação, o agrupamento e dizia, ó, avalei aqui para mim para ver a performance. Tá batendo? Tá.

Opa, beleza. Então, tá validado aqui nessa solução, entendeu? Ou então, eu ia ficar fazendo ciclos de interação mesmo. Então, desenvolver um primeiro modelo, pega um conjunto de 10, 15, 30 candidatos, manda para eles.

Eles verificam lá, me dá o feedback, aí eu pego esse dado, reformulo um outro modelo, gera novamente. Então, faz essa interação, que eu acho também que é bem mais vantajoso, porque o cliente já consegue ver resultado logo de

imediatamente.

Entrevistadora: E aí, você tem uma, é um input multidisciplinar, né? Para fazer essa modelagem, né? Esse direcionamento.

Pedro: É, é muito voltado para o fato da metodologia ágil, né? Então, pega o dado do cliente, desenvolve, entrega para ele, ele dá um feedback, a gente refina e fica nesse ciclo infinito, entendeu?

Entrevistadora: Só voltando uma coisa, se você pudesse me explicar, quando você falou que não usaria a acurácia, porque a base é muito grande, né? Diversa e tal, e aí, o retorno seria um... Você me explica melhor disso, que eu não...

Pedro: A acurácia, ela vê quantos a gente acertou em relação ao total. Vamos pensar nessa base de a pessoa candidata ser aderente ou não à vaga. Então, eu posso ter 100 mil pessoas candidatas, mas apenas sem ser aderente, tá? E aí, o que acontece? Se eu acerto os 999 mil que são não aderentes, eu tô com uma percentual de 99,99

Mas, se eu errei todos que são aderentes, eu não tô fazendo nenhuma... Não tô trazendo valor para o usuário. Então, a acurácia vê quanto que eu acertei, seja positivo ou negativo, se for uma classificação binária, em relação ao todo. Então, se eu errei o que não é... Se eu acertei o que não é aderente, ela vai contabilizar como um acerto para mim, porque eu tô acertando aquilo que não é válido.

Só que eu tô errando todos que são válidos. Então, a acurácia, por si só, ela não é boa por causa disso. Entendi, entendi.

Entrevistadora: E, por fim, assim, eu acho que já é uma coisa que você já introduziu ali, essa parte da efetividade, o impacto na contratação, né? Assim, como que você pode avaliar a efetividade, se realmente estão retornando candidatos que são aderentes à vaga, que vão ser contratados, né? E qual o impacto real na contratação que o seu algoritmo pode ter?

Pedro: Eu acho que efetividade a gente consegue medir, se fosse esse ciclo de interação, porque imagino que os primeiros ciclos pode ser que tenha um número muito grande de não aceito, né? Não, a seleção ali do modelo não seja tão boa, e o cliente vai dizer pô, não tá muito bom. Mas a partir que a gente vai refinando isso, evoluindo esse modelo, a gente tende a trazer mais assertividade, digamos assim, na hora de indicar uma pessoa, uma candidata ali, mais propícia para aquela empresa. Então, acho que ser efetivo seria nisso.

E o impacto, eu imagino que uma vez que a gente consegue validar que tá bom o modelo, vai acelerar bastante. Imagina uma pessoa ter que selecionar ali no LinkedIn, ou qualquer outro portal de vaga, seja lá como for, analisar um por um, e tu tem uma solução semi-automatizada, quase 100

Os processos seletivos em geral são muito longos. Aqui na empresa a fim de início, demora 2, 3 meses para contratar. Então, acho que reduzir isso seria muito bom, né?

Entrevistadora: Sim, legal.

E você acha que, assim, quando você, voltando um pouco na questão da modelagem em si, você acha que se tivesse um framework, uma coisa mais assim, considerando que a gente tá falando de pessoas, né? Você vai garimpar pessoas, e a efetividade disso vai gerar um impacto na contratação, pessoas empregadas ou não. Você acha que se você tivesse um framework para seguir, para talvez melhorar, eliminar a possibilidade de algum viés, alguma coisa assim, seria interessante?

Pedro: Acho que sim. Pensando em viés, óbvio, quando a gente tá tratando de pessoas, aí vem questão de etnia.

Se a gente colocar esse tipo de informação dentro do modelo, cara, já começou errado. E, assim, vai ter viés e viés muito complicados que pode trazer problemas sérios. Então, tendo um framework que conseguiria, mesmo que o usuário, quem tá utilizando, não se atentasse a esses pontos, ele conseguisse já dizer, ó, tem um alerta aqui porque tu colocou o dado que é gênero, etnia, a religião, e esse cara tira aqui dessa análise, porque isso aqui é dado pessoal, não deveria estar envolvido, isso iria ajudar bastante.

Eu acho que iria reduzir muito os problemas que tem hoje por aí.

Entrevistadora: Ah, legal.

Pedro: Mas, no sentido, eu acho que, no sentido de desenvolver por desenvolver, eu acho que pode trazer algumas agilidades, alguma produtividade ali, mas o impacto eu vejo como menos, porque é algo que, se fosse sistema de recomendação, é algo que muita gente já faz no mundo afora.

Então, é o que acaba acontecendo. As pessoas seguem ali já aqueles padrões de desenvolvimento, mas um framework voltado a esses tipos de situações de viés, de ter cuidado com o dado, ter uma qualidade ali, uma curadoria melhor, aí eu acho que é muito, o impacto é maior e não restringe a pessoa querer fazer uma implementação de um outro jeito. Ela pode implementar de A, B, C maneiras, mas a qualidade que ela tem no dado dela é a mesma, independente de qual modelo ela vai usar, entendeu? Então, acho que o impacto nesse ponto, eu acho muito bom.

Entrevistadora: Legal, muito bom, Pedro. Muito obrigada, super rápido você.

Pedro: Nada, que isso.

Foi, foi muito bom, me ajudou bastante essa última pergunta aí, eu já coloquei, assim, já tô mudando a partir das suas respostas, já deu bastante

input legal. Muito bom.

Pedro: Tranquilo.

Entrevistadora: Beleza. Obrigada. Nada, fica tranquila.

Pedro: Qualquer coisinha, é só falar.

Entrevistadora: Tchau, tchau.

B.8

Renato

Entrevistadora: ...se a transcrição tá em português, porque às vezes vai no automático para inglês. Tá tudo certo. Eu vou compartilhar a tela, porque eu acho que fica mais visual.

São só, assim, tópicos para a gente ir guiando a nossa conversa, tá? Mas se você quiser, em qualquer momento, mostrar alguma coisa, sei lá, puxar alguma ferramenta que você queira mostrar, fica à vontade, tá? Que aí você pode pegar aqui o ponto.

Renato: Tá bom.

Entrevistadora: Então, só para a gente iniciar, eu queria pedir para você fazer uma contextualização da sua experiência geral, profissional ou não profissional, às vezes em pesquisa também, com algoritmos em geral, para eu entender, assim, o contexto mesmo.

Renato: Tá. Eu tenho uma dúvida, talvez seja melhor perguntar antes de responder, na verdade.

Entrevistadora: Claro.

Renato: Quando você tá falando de algoritmos aqui, você tá falando algo específico de algoritmos de Machine Learning ou de uma forma mais ampla?

Entrevistadora: É, assim, o que a gente tá fazendo um recorte, porque para minha pesquisa, eu peguei como um exemplo um algoritmo de recomendação, tá? Então, se você faz ou já fez algoritmo de recomendação, é uma maravilha, assim. Mas, de uma maneira geral, o que eu vou pedir aqui para você fazer uma simulação de como você faria um algoritmo de recomendação. Mas, no seu caso, assim, da sua experiência, você pode falar do geral, assim.

Renato: É, eu tô perguntando porque eu não trabalho diretamente com Machine Learning, né? Eu trabalho com desenvolvimento, então tem uma parte ali de engenharia de dados e desenvolvimento de software mesmo. Eu não sei se encaixa no teu, aí, no que você precisa. Por mim, não tem problema.

Se você quiser seguir e analisar depois se vai poder usar esse dado ou não, fica a seu critério. Só para entender que, talvez, a resposta que eu te dê seja até um pouco fora do padrão que você tá esperando, porque eu não trabalho,

assim, eu já fiz uma coisinha ou outra, mas no meu dia-a-dia eu não vou sentar para escrever, para treinar um modelo, alguma coisa desse tipo, entendeu?

Entrevistadora: Entendi. Não, eu acho que, assim, se por você tudo bem, a gente pode seguir, porque, como eu falei, é um exercício até para pensar como você faria, sabe? Se você ficasse com esse desafio.

Renato: Então tá, perfeito. Aí, o meu papel na empresa é, eu tô à frente ali de algumas demandas de desenvolvimento de software para, voltados para a área de ciência de dados, então tem muita interseção, né? Mas são desenvolvimentos de ferramentas que ou utilizam ou, a gente chama de ferramentas habilitadoras, né? Elas dão ferramentas ou fluxos para os cientistas de dados trabalharem. Então a gente acaba tendo muita interação com as ferramentas que os cientistas de dados usam, os fluxos que eles precisam fazer, né? No dia-a-dia, e aí a gente trabalha em cima de ferramentas, né, que dão esse suporte para eles.

E tem uma frente que eu tô trabalhando agora também, que é a de ferramenta generativa, é para a parte de augmented retrieval, né? Que é conversar com bases de dados internas da empresa, sem precisar retreinar o modelo. Então essa é uma parte que já tem um pouco mais de, inclusive, de interseção com ciência de dados, porque tem toda uma forma de esquematizar aquelas coisas ali. E no passado eu trabalhei muito com simulação virtual, então eu já fojo até um pouco disso, trabalhei muito com o desenvolvimento de simuladores mesmo.

Misturava a parte virtual com parte física, voltado para treinamento.

Entrevistadora: Era simulador de que? Só uma curiosidade.

Renato: Teve vários.

Trabalhei com um simulador que é um blindado para o exército, um simulador de navio com a marinha. O da marinha, por exemplo, é para o treinar o prático, né, que é o cara que estaciona o navio no porto, para ele aprender todos os procedimentos para entrar com o navio no porto, sem causar um acidente.

Entrevistadora: Muito legal.

Não, tá ótimo. Então, assim, o que é nessa simulação que eu pedi, né, que é uma objetiva que é entender quais seriam os seus caminhos de pensamento e tomada de decisão, tá? Então, eu vou te passar um contexto, assim, fictício, né, no cenário. É como, assim, imaginar que você trabalhasse numa empresa que faz algoritmos para fazer a divulgação de vagas, né, para conseguir alcançar possíveis candidatos para uma vaga específica.

Daí a gente pode se facilitar para você pensar nas questões técnicas, pode ser, assim, um anúncio no Facebook, um anúncio no LinkedIn, né, não tem

nada muito restrito. E aí uma empresa de tecnologia, de porte médio, contrata a sua empresa, né, contrata você para fazer esse anúncio dessas vagas. E eles listam apenas alguns requisitos, bem, de maneira geral, mas generalista, não dão muitos detalhes.

É uma vaga de tecnologia para um profissional pleno, né, então as skills que precisam, é, precisam convergir para ser um desenvolvedor ou uma desenvolvedora full-stack, com a formação em engenharia de software, tecnologia da informação, sistemas e afins, né, área de tecnologia, de engenharia. É uma experiência média de cinco anos e é para trabalhar em regime remoto. Então, esse é o cenário, são essas as questões que eles passam, né, sem mais nenhum detalhe.

E aí a ideia é pensar assim, isso chega para você, como você faria, começaria esse planejamento e a concepção dessa solução, né, como você define os requisitos e define quais os dados vão ser utilizados, como você escreveria, né, esse, esse modelo, vamos dizer assim.

Renato: Deixa eu tentar entender um pouquinho melhor o problema, a ideia é, tem aquele briefing do episódio anterior, é, mas o resultado esperado é um post, por exemplo, no Facebook, com a divulgação dessa vaga?

Entrevistadora: Isso, você precisa formular um anúncio, né, um post, um anúncio, pode ser até aqueles que ficam aparecendo, né, assim, depende da estratégia, mas assim, é, de alguma forma, algum, esse conteúdo, ele precisa chegar nas pessoas que tenham o perfil que, para ocupar essa vaga, tá? Então, você tá lá no LinkedIn, olhando, aí surge um anúncio de uma vaga, não surgiu à toa para você, é porque você tem, né, o perfil que pode ser, que demete com aquele anúncio, né? Então, a ideia é essa.

Renato: Tá, perfeito, aí pode voltar para as perguntas de novo, por favor.

Os requisitos, né, é, como definiria os requisitos? Tem, o primeiro ponto é entender um pouco disso que eu acabei de perguntar, né, qual era o objetivo, a saída, aí, do, né, do, esse algoritmo, né, o que que ele vai devolver, é, também entender, é, como eu estava falando de recomendação, não é, como eu falei, não é minha praia, mas eu imagino em alto nível aqui que a gente quer fazer alguma coisa mais direcionada, então, provavelmente, esse algoritmo, essa, ele não vai gerar um, uma, um post e simplesmente jogar para uma lista de candidatos, provavelmente, seria mais eficiente gerar variações desse post, com, de acordo com os perfis dos usuários que vão receber, né, do, lá do, no caso do LinkedIn, né? Então, é, e aí tentar fazer algum levantamento com o cliente, é, ou com conhecimento da empresa, também de histórico, de quais seriam as melhores formas de fazer essa segregação, né? Se, é, tem ali, né, no requisito de ser, acho que perfil pleno, tempo de experiência, mas aí, se, por exemplo, se não tem uma

exigência de ser uma linguagem específica, talvez esse, esse, possa ser gerado de postos diferentes, com, dando mais destaque para uma linguagem ou outra, e aí, ele vai pedir, por exemplo, vai colocar como exemplo que uma linguagem aceita a S-Sharp e vai mandar isso para o desenvolvedor S-Sharp, é, e isso, em termos de tecnologia, né, pode ser qualquer outra tecnologia que esteja ali no meio, para, para aquela vaga se tornar mais atraente, né, para aquele, para aquele usuário que vai receber, não ficar algo muito amplo, é, talvez no tempo de experiência, alguma coisa semelhante, porque ali, né, acho que estava escrito cinco anos, mas provavelmente dá para trabalhar dentro de uma margem, né, então, é, fazer alguma recomendação assim, e aí, já conectando na parte de baixo, né, assim, dados utilizados, é, né, teria, eu acho que teria essas duas frentes, né, o dado do lado do cliente, o que que faz sentido direcionar o, direcionar o, o processamento, né, para essas classificações, digamos assim, e do lado do, de quem vai receber, aí, os dados que forem possíveis. Então, assim, no caso do LinkedIn, como está acessando um currículo, então, é, daria para pegar o tempo de experiência, talvez, é, priorizar, não, não sei se eu utilizaria isso, mas algo que poderia ser colocado no algoritmo seria priorizar se a pessoa tem experiência em algumas empresas, por exemplo, empresas grandes de tecnologia, ou se ela tem formação em universidades, é, em um conjunto de universidades mais conhecido. Eu não gosto muito dessa abordagem, porque acaba excluindo, né, talvez um bom candidato, só porque ele não se formou em um lugar A ou B, mas seria algo que poderia, né, dependendo do contexto ser usado, é, e outras informações do currículo dele sobre, é, a experiência.

Então, lá no, usando o LinkedIn mesmo como exemplo, pegar ali, é, tecnologias e outras competências que o próprio usuário consegue listar no perfil dele, ou, é, recomendações que outros usuários dão dentro do perfil, né, daquela pessoa. É, hoje em dia, com questões de LGPD, também já, provavelmente, coisas muito mais pessoais, talvez não fosse muito adequado colocar aí no algoritmo, né, por exemplo, idade, é, ou, local onde a pessoa mora, talvez, dependendo do tipo de, acho que estava ali, né, não lembro.

Entrevistadora: Era remoto.

Renato: Era remoto, então, o local onde ela trabalha, talvez ainda seja relevante, porque, remoto, se você for contratar como CLT no Brasil, talvez a pessoa tenha que, pelo menos, morar no Brasil para fazer alguns processos, mas, fora de, a cidade, o estado, talvez já não seja importante. É, e aí, outras informações, acho que não tem nenhuma outra informação pessoal que seria, realmente, necessária ali na hora de, de fazer essa recomendação, né, acho que seria mais, realmente, a localidade para essa questão de garantir que a pessoa pode trabalhar naquele local e, e não colocaria outras informações como idade,

gênero, esse tipo de coisa que, provavelmente, não seria muito adequado colocar nesse tipo de recomendação, se estiver tentando pensar num caso ideal, né, porque, provavelmente, algum gestor pode querer alguma coisa do tipo, ah, eu não quero pessoas acima de uma idade, mas aí já está introduzindo ali no algoritmo um preconceito ali que pode não ser adequado para esse processo.

Entrevistadora: Até, Renato, se você quiser, também, assim, eu acho que você já colocou bastante, mas pensando com a sua cabeça de quem contrata, né, também, a sua experiência, porque você tem, você faz, né, entrevistas, procura pessoas, né?

Renato: É, porque quando eu estou abrindo uma vaga, aí, assim, no caso, é uma oferta só, né, acaba não tendo nenhuma personalização nesse sentido e aí eu não faço a parte de ir atrás dos candidatos, mas aí os dados na vaga são bem objetivos, né, um range de experiência ou um mínimo quando, ah, eu preciso de uma pessoa com, pelo menos, tantos anos de experiência, e, e, e tecnologias que, e, tecnologias e, e, né, tipos de, de atuação, assim, ah, você é um desenvolvedor front-end, um back-end, alguma coisa do tipo, eu como, o que que ela vai fazer, né, para ela garantir que ela tem experiência naquele tipo de trabalho, que tenha conhecimento de uma variação ali, algumas possibilidades de tecnologia, geralmente, mesmo que eu precise de um cara, de uma tecnologia específica, eu não costumo fechar para aquela tecnologia, se ele conhecer algo semelhante, ainda vale a pena entrevistar e ver se aquela pessoa consegue, geralmente, um cara bom consegue se virar com uma tecnologia e, às vezes, se ele é, ah, eu só sei mexer nisso, às vezes, nem é um bom sinal, né? Então, sempre dou um pouco de flexibilidade, assim, na hora de montar a vaga para ter um range maior de candidatos mesmo e, e também para a vaga ser atraente para o candidato, né, para ele não ver de cara, ah, aqui pede, é obrigatório ter C Sharp, eu não conheço tanto C Sharp, então nem vou me candidatar.

E aí a gente costuma pedir também, né, no que eu estava falando de formação e tal, não peço alguma formação específica numa universidade específica ou algo do tipo, né, mas ela tem que ser formada nessas áreas aqui, porque aí garante que a pessoa vai ter uma base do que a gente precisa, mas aí o foco está muito mais em garantir que ela tem uma formação mínima do que ela ter vindo da Universidade de A ou B, né?

Entrevistadora: Entendi, bom, é legal. Então, é assim, pensando, né, nessa, passando essa fase inicial das definições, você faria, como você faria para fazer algum teste necessário, né, para realizar, para validar o algoritmo e que métricas de sucesso você adotaria, né? Lembrando que, assim, você foi uma empresa contratada para entregar um resultado, né, então com certeza vai ser

exigido que tenha alguma métrica de sucesso para você validar a sua solução, né? E se você considera algum feedback para essas iterações, né, porque a gente sabe que vai testar, vai rodar, né, e vai ajustando, então esse é o contexto dessa fase.

Renato: É, em termos de testes, acho que a cola que está aí, ela talvez esteja me direcionando, mas, mas sim, é fazer testes A, B, né, dentro do, dentro de um mesmo, considerando, né, que a gente está gerando vários anúncios, mas mesmo, para o mesmo público-alvo ter variações ali e para tentar medir, né, qual pode ter maior eficiência.

Em termos de métricas, uma provavelmente seria o número de candidaturas, mas talvez também de cliques ali, de rastrear os cliques na vaga, porque não necessariamente, ou até de pay-to-view mesmo, né, de visualizações. O cara está só passando, mas chegou para ele, ele chegou a visualizar. Então, né, resumindo, visualizações, interações, né, para ver mais detalhes e tal, e de fato o candidato se inscrever na vaga.

É, e em termos de feedbacks, é, talvez cruzar, né, os anúncios, né, as vagas que foram, as vagas que foram geradas com os candidatos que realmente se candidataram naquela vaga, né, para ver, por exemplo, que eu tenho uma vaga de pleno e eu estou recebendo muitos candidatos júnior, tem boa chance de ter algum problema na descrição da vaga, né. Claro que alguns vão se candidatar, né, porque, ah, eu vou me arriscar aqui, que de repente eu consigo. Ok, mas se o número está muito grande, provavelmente os requisitos não estão bem claros ali na vaga, ou está chegando gente com perfil muito fora do que está especificado na vaga.

Também seria uma forma de tentar avaliar se as vagas estão sendo enviadas para as pessoas corretas, né. E talvez também fosse possível, aí é mais uma questão da plataforma, não sei se é possível pegar essa informação lá ou não, mas vamos assumir que sim. Por exemplo, se a gente conseguir visualizar, pegar para quem o anúncio foi entregue, talvez desse para olhar ali, né, ah, esse anúncio foi entregue para essa lista de pessoas e aí também fazer uma análise ali se ele foi entregue, se fez sentido ou não as pessoas para quem ele entregou.

Entra um pouco na métrica, mas seria também uma forma de identificar, não, ele está entregando muito para o público A, mas ele foi produzido para ir para o público B. Então, alguma coisa na lógica do algoritmo não está correta e ele precisa ser ajustado ali para realmente ir para o público que deveria.

Entrevistadora: Tá, legal. E aí, assim, por fim, pensar se olhar para esses resultados e mensurar o impacto na contratação, que tem algum retorno que faria você pensar diferente, né, assim, se você precisa pensar, você já

falou assim, né, pelas métricas, você vai querer saber quantos que realmente convergiram, né, quantos que retornaram e tudo mais, mas pensando no impacto da contratação em si, né, já pensando no contexto mais localizado na sociedade, você acha, né, que para você conseguir mensurar, entender esse impacto, você precisaria trocar com alguém do negócio em si, você colocaria isso como uma solução, né, para você pensar, chamar alguém de contratação, alguém de RH para trocar uma ideia ou você ficaria mais na troca mesmo com seus pares de tecnologia, como seria essa abordagem, né, para fazer essa calibração, vamos dizer, entender o impacto do seu algoritmo?

Renato: Não, eu acho que seria importante porque, geralmente, nos processos, tem vários níveis e a parte técnica é só um deles, né, então, alguém mais especializado em avaliar uma parte comportamental, de soft skills, por exemplo, sabe que a gente olha um pouco, né, pensando aí na posição do entrevistador, mas aqui na Deloitte mesmo tem a área do talent que vai fazer uma entrevista voltada para isso, então, essa parte seria importante entrar ali também, de tanto na avaliação lá no final, mas também até talvez lá no começo em ajudar a definir algum, é, algum, né, tanto naqueles parâmetros, né, definir alguma coisa do tipo, é mais difícil porque é mais subjetiva, mas é ter alguma coisa do tipo, será, se a pessoa trabalha bem em equipe, se é comunicativa, para a recomendação também seria difícil o algoritmo chegar nessa conclusão para usar isso como um parâmetro, mas é algo que seria importante já ter mapeado desde o início para saber, ah, essa é uma vaga que a pessoa vai trabalhar muito em equipe, então, no processo seletivo isso tem que ser avaliado e colocado ali, né.

Então, acho que sim, é, outras áreas, uma área de RH, assim, de, seria importante para ajudar a avaliar, é, a eficiência do algoritmo sim.

Entrevistadora: Tá ótimo, tá legal. Acabou, a ideia era mesmo pensar assim, né, investigar como seria seu raciocínio, não sei se você quer colocar mais alguma observação específica em relação a esse processo, alguma experiência, alguma coisa?

Renato: Não, acho que não, mas interessante, interessante a pesquisa aí, espero ter resposta aí, ajudem, né, o trabalho.

Entrevistadora: Com certeza, não, é ótimo, principalmente, assim, eu tô entrevistando algumas pessoas que não têm esse, essa experiência específica com recomendação e eu acho que, assim, é até, tá sendo até mais positivo, sabe, porque você vai mais aberto para pensar nas possibilidades, né, muito bom, legal.

Renato: Perfeito.

Entrevistadora: É só te falar, eu não falei no começo porque eu ia

atrapalhar, eu trabalhei há um tempo, não sei se você conhece o laboratório da Coppe, de, que o pessoal fazia simulação de, para treinamento de avião e de guindaste?

Renato: Ah, é, eu já fui lá, esqueci o nome dele agora.

Entrevistadora: É o Lance? Lance, é, eu conheci lá. Conheceu? É. Ah, eu trabalhei 10 anos lá, junto com o Gerson, com o Landau, aí você foi falando isso, por isso que eu falei assim, gente, a gente, porque acaba sendo meio nichado, né, essa parte de simulação.

Renato: É, assim. É, legal. Esse Lance é bem legal, né, tem um prédio bacana lá, né?

Entrevistadora: Tem, lá no fundão, eu trabalhava naquele prédio. Legal.

É lá, né, bem longe, né, o dia a dia não era tranquilo, não. É, para chegar. É, mas era legal, a gente brincava de tecnologia lá.

Renato: É, esses projetos de simulação, assim, eu gosto bastante, porque sempre tem, aí, às vezes, quando eu trabalhei na marinha, tinha muita integração com o hardware, né, então, às vezes, o cara chegava lá com o hardware maluco que ele tirou do navio, ah, vamos tentar botar isso no simulador, aí tem que fazer uma engenharia reversa para entender como o troço funciona, bem, era...

Entrevistadora: Ah, maravilhoso.

Renato: Bem, cientista maluco mesmo, de ficar funcionando as coisas e tal, divertido. É, bem isso, a gente chegou a construir uma cabine fake, né, assim mesmo, porque eu sou designer, e aí eu trabalhava com a parte de produção do design, de produto mesmo, né, com o pessoal.

Entrevistadora: A gente construiu uma cabine, aí, depois a gente ganhou até, eu não lembro se foi da aeronáutica, eu não sei, a gente ganhou um domo gigantesco que o pessoal usava para simular, e aí a gente fez um monte de invenção lá, era legal, era cientista maluco mesmo. Então, tá bom, Renato, muito obrigada, viu?

Renato: Nada.

Entrevistadora: Bom fim de semana para você, tudo de bom.

Renato: Para você também. Tchau.

Entrevistadora: Tchau, tchau.

B.9

Roberto

Entrevistadora: A descrição do Teams não tá tão horrível não, tá, em português. Tá ótimo.

Roberto: Ela é boa, ela é boa.

Entrevistadora: É, eu vou compartilhar o material aqui com você, mas são só, assim, direcionadores para nossa conversa, tá? A ideia é uma entrevista semi-estruturada, mas a gente usa um método de design, que é o Directed Storytelling, que é, eu vou pedir para, vou te dar alguns inputs, te passar um cenário fictício, e aí a ideia é você ir trazendo tudo que vem na sua cabeça, mesmo que não esteja ainda estruturado, assim, só que eu quero analisar, né, assim, tem as questões técnicas, mas eu quero analisar de verdade o seu raciocínio e os seus caminhos de tomada de decisão, tá? Então, essa é mais importante, assim, pode ir falando, mesmo que depois você fale, não, melhor outra coisa e tal, sem compromisso com essa coisa muito rígida, tá? E aí, se você quiser compartilhar, não sei, às vezes, se você quiser mostrar alguma tela ou alguma ferramenta, alguma coisa, fica à vontade. Tá bom, obrigada. Para começar, eu queria só uma contextualização, assim, para você, se você puder me contar um pouco sobre a experiência que você tem desenvolvendo algoritmos na sua vida, né, profissional, e um pouco do seu papel no seu trabalho hoje, né?

Roberto: Bom, eu comecei computação em 2005, né, até 2010, a partir de 2010 a 2016 eu fiz mestrado e doutorado, né, então, assim, a minha experiência de criação de algoritmos foi praticamente toda a computação baseada em algoritmos voltados para pesquisa, né, para parte de engenharia de software, desenvolvimento de sistemas multiagentes, né, seria como se fosse uma inteligência artificial distribuída, e posteriormente comecei a trabalhar no mercado, então, tive alguma experiência aí com o desenvolvimento de algoritmos para as situações aí do dia a dia, de óleo e gás, também laboratórios de inovação na Petronet, e hoje em dia eu trabalho com o desenvolvimento de soluções AWS, e também tem um projeto que a gente trabalha com as ferramentas da Microsoft, que é o Power BI, o SharePoint, o Power Apps, que são baseados em low-code, né, assim, você deixa um pouco a codificação de lado e trabalha mais ali, como se fosse arrastando caixinhas, né, e organizando essas caixinhas ali para poder mostrar a parte analítica dos dados do cliente.

Entrevistadora: E quando você trabalhou com, você estava falando de A distribuída, né, quando você trabalhou também nessa parte de pesquisa, você desenvolvia mais algoritmo de classificação, tinha algum tipo específico assim, ou eram diversos?

Roberto: Eram diversos, assim, nós trabalhávamos com algoritmos, por exemplo, algoritmos normativos, a ideia do algoritmo normativo na inteligência artificial é a seguinte, a inteligência artificial, ela não segue regras, então ela vai tentando ali, né, achar o caminho dela para resolver o problema, mas esses algoritmos, eles não podiam estar no sistema reveria, eles tinham que ser

organizados de alguma forma, você não podia restringir eles, mas você também não podia deixar eles totalmente autônomos, então a gente criava normas ali no sistema para a inteligência artificial, ela verificar se ela ia poder solucionar ou não um determinado problema, dado uma norma, essa norma podia permitir essa inteligência artificial solucionar esse problema ou podia restringir, então às vezes a inteligência artificial, ela acabava resolvendo um problema que ela não era o papel dela resolver, mesmo ela recebendo uma punição, porque posteriormente os algoritmos iam adquirindo inteligência e iam entendendo que mais para frente eles iam ganhar mais recompensas, mesmo resolvendo problemas que não eram deles, então a gente trabalhou bem em cima disso. Um pouco de trust, reputation, que é a parte de você qualificar ali a reputação a algoritmos de inteligência artificial que você não conhece, por exemplo você tem um leilão de quadros, você quer comprar um quadro de um determinado pintor e aí você não conhece esse pintor, mas aí você começa a perguntar para outros algoritmos que estão naquele leilão, outras inteligências artificiais, e aí essas inteligências artificiais começam a se comunicar para verificar qual que é a reputação dos vendedores e dos compradores para poder te dar um retorno com relação a fechar negócio ou não com aquele vendedor.

Entrevistadora: Interessante isso, bem legal.

Roberto: A gente trabalhou bem nesse nível de pesquisa.

Entrevistadora: A gente começa aqui nossa simulação, vou te passar um contexto, um objetivo, um cenário. Imagina que uma empresa de tecnologia de porte médio contrata outra empresa para fazer uma divulgação de vagas, para anunciar uma oportunidade, né? E você é o contratado, você está nessa empresa que vai fazer o algoritmo para rodar um anúncio com uma oportunidade e eles só trazem esses requisitos, que é uma vaga de tecnologia para profissional pleno, com skills que atendam para ser um desenvolvedor, uma desenvolvedora full stack, com formação em engenharia de software, tecnologia da informação, sistemas e afins de tecnologia, engenharia, mais ou menos cinco anos de experiência e vai ser regime remoto.

Então eles te contratam, através da empresa e falam desenvolve um anúncio para rodar em alguma plataforma e a gente não precisa se ater a plataforma específica, mas assim só para deixar mais concreto, né? Imagina aqueles anúncios que aparecem no Facebook quando você está navegando, né? Aparece algum até próprio Instagram, mas assim vaga de anúncio não aparece muito no Instagram, mas alguma coisa que te capte para entrar nesse anúncio e você como candidato se candidatar àquela vaga, tá bom? Daí a primeira parte é a gente pensar assim, como você faria esse planejamento e a concepção dessa solução? Pensando na definição dos requisitos, né? Como que você define

esses requisitos, como você, assim, se precisar de mais informação, como você busca deixá-los mais concretos, né? E mais destrinchados ali e quais dados você utilizaria, né? Pensando assim, o que você pode buscar para usar como dado para sua solução.

Roberto: Tá bom. Com relação à definição dos requisitos, existe algumas formas de você extrair isso ali de uma conversa com o cliente, né? Então assim, você pode utilizar a entrevista, que é algo assim que você senta, né? E faz perguntas não estruturadas ou semiestruturadas para a pessoa entender.

Você primeiro tem que entender o problema dela, verificar se aquilo de fato é um problema que dá para você gerar, né, um software ou algo que vai solucionar o problema dessa pessoa. A partir desse ponto, você vai começar a direcionar essa pessoa de forma que ela consiga mostrar para você o que ela precisa em forma de funcionalidades. Então assim, nós temos, por exemplo, a possibilidade de sentar para a pessoa e falar assim, olha, o que você imagina nesse seu sistema? E aí você deixa a pessoa falar, ela explicar o que ela quer, o que ela gostaria.

Provavelmente ela já viu alguma coisa parecida antes na internet, vai te trazer ali alguns sistemas próximos ao que ela imagina, e você vai poder verificar com ela na entrevista qual que é a finalidade. Assim, na extração de requisitos, é o momento em que você mais escuta o seu cliente. Nós podemos fazer isso também na forma de, eles falam que tem uma técnica chamada saco de requisitos, que as pessoas vão escrevendo ali tudo que elas acham que deve ter no sistema, coloca isso numa sacola, e à medida que as pessoas vão tirando aquilo ali, vocês vai debatendo aquela real necessidade ali com o cliente, e para definir as funcionalidades do sistema.

A que eu mais tenho utilizado hoje é a reunião com o cliente, seguido de uma pessoa desenhando junto com o cliente, e anotando ali o caminho principal dos requisitos com o cliente. Então essa para mim tem se demonstrado ser a mais eficaz a nível de mercado. Por exemplo, com relação à parte de dados, feito esse levantamento dos requisitos, você tem a possibilidade de criar ali uma primeira versão rápida para mostrar para o seu cliente, para validar se está fazendo sentido a mensagem que ele te passou.

Então você pode usar softwares que você não precisa programar, apenas você fazer o desenho ali de telas, para mostrar para eles um fluxo do sistema. E usar dados, geralmente dados de bancos abertos. Uma das maiores dificuldades que a gente tem na hora de demonstrar para o cliente rapidamente um exemplo do que vai ser feito, do que foi feito, é a utilização de dados.

Por exemplo, o seu cliente geralmente vai dificultar ao máximo de passar os dados, porque ele acredita que aquele dado dele é ouro, e você pode fazer mau

uso dele. Só que ao mesmo tempo ele está prejudicando a criação do software ali, aquela primeira visão, porque quando você mostra para ele com dados que você achou na internet, com dados fakes, com dados que você cadastrou ali, a primeira coisa que um cliente fala é assim, mas esse dado está errado, esse dado não faz sentido. Então é bem interessante, esse é o primeiro contato que eu tenho com todo cliente do mercado, que eles estão preocupados, eles não estão preocupados com a funcionalidade, com o caminho ali que o software tem que levar de ponta a ponta, ou talvez eles não entenderam ainda que esse caminho é mais importante do que o dado que foi colocado ali como uma primeira amostra, sabe? Então, isso aí nessa primeira entrega ali que seria, por exemplo, de um MVP de telas ali, isso eu vejo como uma dificuldade grande de mostrar para o cliente, assim, é um resultado.

Entrevistadora: Tá, mas assim, pensando que você deveria construir um perfil, né? Assim, você vai fazer um algoritmo que vai, pode ser um algoritmo de recomendação, que a partir do momento que você tiver definido, né, o anúncio, o que deve ter naquele anúncio, você vai procurar pessoas, o anúncio, né, vai procurar pessoas que tenham fit com esse, com esses requisitos, né, que vocês definiram. Você tem uma ideia de quais dados seriam importantes para você entender se a pessoa cabe nesses requisitos?

Roberto: Sim. Primeira, não seria um preconceito, mas assim, a pessoa ter uma graduação, porque, assim, as pessoas que trabalham na área de tecnologia e não possuem graduação, elas têm muita dificuldade para demonstrar o que elas sabem de fato, ou que pelo menos elas têm a base necessária para aprender novas tecnologias, então isso eu vejo como uma dificuldade, então assim, esse seria o primeiro ponto, tá? Com relação à parte de desenvolvimento, se essa pessoa colocou no currículo que ela sabe as novas tecnologias que estão no mercado e se ele faz uso delas, e se ele tem alguma forma de demonstrar que ele tem esse conhecimento, por exemplo, que a pessoa tem um git, né, onde você possa ir lá e ver o código que ele desenvolveu, então assim, isso é muito bom, porque você vai lá, mexe, vê o que o cara já fez, você roda no seu computador e fala assim, legal, o cara possui esse, esse e esse conhecimento, já dá para você saber se realmente ele é o full stack, né, se ele consegue criar ali do início ao fim uma determinada aplicação, ou se ele está te passando ali no currículo apenas uma ideia geral do que ele já passou o olho, sabe, no mercado, tá? A experiência de cinco anos, cara, eu não vejo ela como uma necessidade, assim, apesar de ser um requisito aí do cenário, porque assim, as tecnologias mudam muito rápido, então assim, dificilmente você vai ter uma pessoa que tenha cinco anos em um determinado ramo de tecnologia, você vai pegar aí uma pessoa que vai estar dois anos na empresa, vai migrar mais um para outra,

mais dois para outra, então assim, talvez a pessoa saiba um pouco de Python, um pouco de AWS, um pouco de, um pouco de Java, nessa parte eu acho que assim, o requisito diferencial, se você tivesse como aplicar um desafio para essa pessoa, assim, anunciar que vai ter um desafio ali para essa pessoa desenvolver, é mais interessante do que você validar todas essas características, assim, só na entrada do sistema, sabe?

Entrevistadora: Entendi. Tá bom.

Roberto: Respondeu?

Entrevistadora: Respondeu, respondeu sim. E aí, sobre o desenvolvimento do algoritmo em si, você consegue pensar, assim, qual método você poderia escolher, se tem algum desafio técnico que te faz até escolher determinado método ou não, e aí, assim, mais nessa parte mais técnica mesmo.

Roberto: Olha, para gerar o anúncio, né?

Entrevistadora: Isso.

Roberto: Então, para gerar o anúncio, você poderia pegar um algoritmo de comparação, você gera um algoritmo ali onde você tem uma base, né? Com os currículos já cadastrados, você roda, você pega, por exemplo, 30

Então, ele vai tender a te trazer os currículos que você espera, que tem aquelas características, né, que você gostaria, dado ao que você já calibrou ele em cima de 30

Entrevistadora: E aí, para você dizer, assim, quais são, por exemplo, você pegar um currículo como modelo e dizer para o algoritmo, né, isso aqui é o ideal, você teria alguma técnica, assim, você vai bater com os requisitos, vai ser uma inferência sua, como seria?

Roberto: Então, inicialmente poderia ser uma inferência nossa para você fazer uma linha base, tá? Por exemplo, assim, você pega um algoritmo de machine learning, não precisaria de ser, porque assim, tem algoritmos que seria para você pegar um canhão para matar uma formiga, mas você poderia pegar um algoritmo ali de comparação de atributos e montar o seu perfil, você poderia usar uma rede neural, você poderia usar um K-means, né, que faz o que? Ele pega esse K, que seria a ponderação ali do algoritmo, e ele vai fazendo um balanceamento entre os atributos ali que você passou da vaga para verificar a pessoa que cobre, de certa forma, aqueles atributos.

Então, assim, mesmo que a pessoa tenha um nível baixo em um atributo, ela tem um muito alto em outro, que talvez seja mais o perfil da vaga que você quer preencher. Então, seria um algoritmo bom de ser utilizado.

Entrevistadora: Legal. Mas aí você vai ajustando de acordo com o que você pegou ali de input inicial, né? Ou você, nesse momento, tem participação

também, alguma maneira de ter participação do cliente, alguma coisa assim, é mais técnico?

Roberto: Sim, não, você consegue colocar participação do cliente porque você consegue modelar um V0 ali, um ponto de partida, com características que o cliente acredita ser interessante para aquela vaga.

Entrevistadora: Tá.

Roberto: Você consegue fazer isso, sim.

Entrevistadora: E aí, pensando que assim, você foi contratado, né? Você foi contratado, então você tem que mostrar, assim, de repente, pensar em algum teste, né, para saber se está funcionando, se está sendo ajustado ali de maneira assertiva e levantar alguma métrica de sucesso, né, para mostrar de volta para o seu cliente e pensar em algum, não sei, né, se você pensaria em feedbacks para as interações, né, feedbacks humanos, não humanos, para ir fazendo a melhoria dos resultados.

Roberto: Tá. Começando pelos tipos de teste, você poderia fazer testes, por exemplo, de caixa preta, que são testes que você insere dados que, por exemplo, o cliente não sabe quais são ali, você não tem a visão do que está dentro dessa caixa, por isso que chama caixa preta, né? São dados que você insere para verificar se o algoritmo consegue excluir ou incluir dentro da escolha dele.

Então, assim, esse é um teste bom para você fazer com inteligência artificial. Você pode também passar testes caixa branca, que são testes que você passa as variáveis para o algoritmo decidir isso. Então, você deixa as variáveis claras ali, que poderiam ser definidas, por exemplo, com o cliente, o que ele imaginaria que seria ali um pequeno peso para cada característica da pessoa que ele gostaria de contratar, inserir isso no sistema e verificar se o algoritmo vai pegar ou não aquele dado para você.

Você consegue fazer isso também. E você pode também fazer essa calibragem inicial, você pode pegar essa calibragem também e ir modificando ela ao longo do teste para verificar se alguma coisa ficou de fora do que você esperava no algoritmo. Isso também é possível de fazer.

Você vai mudando a calibragem do que você espera de saída, o peso das variáveis que você acredita ser importante para a vaga e, naturalmente, ele vai ter que alterar os resultados. Se ele não alterar esses resultados, o algoritmo não está com a precisão que você imagina para a curácea, que acho que é o termo, de achar o dado que você gostaria. Métricas que seriam utilizadas, a curácea, de voltar os dados que você tem uma noção de quais são, você pode colocar também como métrica que você teria que fazer inicialmente, dada a inserção de testes que você já conhece, tipos de currículos que você já conhece, quais

entrariam não em métricas de sucesso ou de erro, e se não foram encontrados, por quê? Então, se o seu teste, de certa forma, está falho, se aquela métrica não apresentou o dado que você esperava que tivesse ali.

E com relação a feedbacks e interações, como avaliar retornos, etc. Esse feedback e interações seria em qual sentido?

Entrevistadora: Você até já falou um pouco, mas assim, por exemplo, se está tendo um retorno positivo, e aí vem um pouco da métrica de sucesso, tem muita gente se inscrevendo na vaga, ou o feedback da própria pessoa que te contratou. Está vindo muita gente que não tem nada a ver, enfim, te dar um feedback ali.

Roberto: Isso é bem interessante, porque você pode avaliar isso, dadas as características que você colocou no anúncio. Porque, por exemplo, cinco anos de experiência, isso eliminaria muita gente boa, e pegaria caras que talvez estivessem obsoletos, que estivessem trabalhando há cinco anos, dez anos na Petrobras, com as mesmas tecnologias, e não conhecesse nada de novo que está rodando no mercado. Então, assim, isso eu acho que seria um ponto de observação, por exemplo.

Ou você falar que o cara precisa ser full stack. Quando você fala isso, você está falando assim, eu quero um cara sênior, mas a vaga foi anunciada para pleno. Então, um cara que é pleno, ele vai ter vários buracos ali na formação dele ainda.

Ele já não tenta de qualquer maneira ir para um júnior, mas ele ainda não tem o conhecimento necessário para sentar e realizar aquela tarefa ali com o tempo que ele estipulou. Com relação a feedbacks para o algoritmo, o algoritmo podia encontrar isso. O algoritmo, dependendo da forma com que você colocar o peso nos atributos, ele vai te trazer perfis diferentes.

Então, muitas vezes a pessoa vai responder que ela preenche todos os requisitos. E o algoritmo vai te trazer sempre essas pessoas, porque elas responderam o que estava com peso ali. Então, o que você esperava que fosse selecionado inicial.

Uma coisa que poderia melhorar isso talvez fosse a capacidade da pessoa especificar quais são as tecnologias que ela utiliza na hora de falar da parte de potencial de programação e desenvolvimento da pessoa. Então, esses detalhes trariam mais assertividade ao algoritmo.

Entrevistadora: Por fim, acho que você já trouxe um pouco, mas pensar qual seria a efetividade mesmo, se você conseguia avaliar a efetividade do algoritmo para os candidatos em relação às pesquisas desejadas.

Acho que foi isso um pouco que você já falou. E o impacto na contratação como um todo, pensando qual é o efeito que o direcionamento do seu algoritmo

pode causar na contratação de uma maneira de sociedade, uma coisa mais ampla mesmo.

Roberto: O algoritmo pode ter um grau de efetividade quando você sabe direcionar a vaga.

Então, você tem que ajudar o algoritmo inicialmente a conseguir filtrar os primeiros exemplos de currículos que você deseja. Então, ele vai conseguir fazer isso. Impacto na contratação.

O algoritmo fica dependente de entradas. Então, dependendo da forma como a pessoa coloca o currículo dela, apresenta as características, ele pode ou não ser selecionado. Talvez um cara muito bom que teve preguiça e escreveu o currículo dele de uma forma correta, ele talvez não fosse chamado.

E um cara que soube utilizar os atributos que estavam como pré-requisitos na vaga e os caracterizou bem, o algoritmo poderia escolher essa pessoa. Então, existe essa complexidade no impacto da contratação. Ele poderia gerar esse erro.

Entrevistadora: Tá. Oi.

Roberto: Estou te ouvindo. Eu até fechei a câmera aqui para ver se melhora um pouquinho.

Entrevistadora: Eu acho que mesmo o nosso espelho aqui na nossa entrevista, que ele é mesmo a maioria. Conseguiu falar bem assim.

Roberto: Tá picotando bem a sua voz. Fechei aqui também para ver se melhora. Se melhora, né? Então, eu gostei muito da entrevista.

Achei bem interessante a forma como você trouxe. Você trouxe o projeto e a implementação de software. Então, você passa por todas as partes, deixando bem claro que uma má criação do levantamento dos requisitos iniciais vai acabar por comprometer ou trazer ali um desenvolvimento de software adequado para a necessidade do cliente.

Entrevistadora: Muito bom. Vou citar aqui só mais uma pergunta. A gente está falando de petô, os petôs vão funcionar.

Você acha que se tivesse um prêmio para melhorar, melhorar não, vamos dizer assim, direcionar questões para mitigar...

Roberto: Larissa, eu não estou te ouvindo. Eu ainda não estou te ouvindo.

Entrevistadora: Por quê?

Roberto: É... Você pode repetir? Eu não consegui ouvir.

Entrevistadora: Você está me ouvindo agora?

Roberto: Estou ouvindo, mas está falhando bem. Será que é minha internet? Depois eu tento ficar mais perto do moldem. Você consegue me ouvir?

Entrevistadora: Sim, eu estou ouvindo. A internet... Sim.

Roberto: A sua voz está bem picada. Você falou da questão de criação de um framework? Oi? Você falou da criação de um framework? Sim. Está bem. Deixa eu tentar. Deixa eu ver se a minha internet aqui...

Entrevistadora: Acho que agora troquei aqui. Vê se melhorou.

Roberto: Melhorou. Bem. Agora eu estou conseguindo te ouvir. Você está me ouvindo? Vamos ver se você está me ouvindo. Aí você me fala.

Entrevistadora: Estou ouvindo. Está ouvindo?

Roberto: Estou ouvindo. Você me explica de novo a parte do framework?

Entrevistadora: Sim. Sim. Você acha... A gente está falando que vão precisar... ...contratação, enfim, até uma questão mais ampla com outros recortes. Você acha que se tivesse um framework para te auxiliar a fazer a mitigação de vieses, seria interessante logo nesse início da modelagem?

Roberto: Sim. Larissa, a resposta é sim. Primeiro porque você evita abrir muito o leque de possibilidades. Porque você vai... Você sendo direcionado, você não precisa pegar diversos currículos nesse problema que não teriam a ver com a sua necessidade.

Depois você traria um facilitador que, dado o framework, você poderia estender os pontos de conexão desse framework para auxiliar na solução do problema. Ele te daria o caminho dado que você passou para ele as suas necessidades e ele te retornaria o resultado. Eu acho que isso é um ponto bem interessante.

Entrevistadora: Bom, muito obrigada. Falhas técnicas, mas muito.

Roberto: Não escutei o final, desculpa.

Entrevistadora: Eu agradei. Desculpa as falhas técnicas, mas pode...

Roberto: Que bom, cara. Espero que possa te ajudar aí, tá bom? E se você precisar de mais alguma coisa, eu fico à disposição.

Entrevistadora: Muito obrigada mesmo. Tá bom. Obrigado aí.

Roberto: Um bom abraço, viu? Vai estar tudo certo lá. Fecha isso aí logo.

Entrevistadora: Tomara, tomara.

Roberto: Tá jóia. Muito bom. Um abraço aí, boa noite, Larissa.

Entrevistadora: Muito abraço.

B.10

Vinicius

Entrevistadora: Porque, às vezes, ela volta automático para o inglês aqui.

Vinicius: Sim.

Entrevistadora: Foi.

Então, eu vou compartilhar a minha tela, é só, são, assim, uns direcionadores, tá, para a gente se guiar na nossa conversa aqui, mas, se você quiser

pegar a tela para projetar, sei lá, mostrar alguma plataforma, alguma coisa que venha à sua cabeça, fica à vontade, tá?

Vinicius: Pode deixar.

Entrevistadora: E aí, assim, o começo, eu queria que você, se você puder, dar uma contextualização na parte profissional, assim, né, que você tem, na experiência, principalmente com algoritmos, né, alguns projetos que você já deve ter desenvolvido, não precisa entrar no contexto dos projetos, né, da temática, mas, de uma maneira geral, dar um panorama, assim, para a gente entender qual é a experiência, o aprofundamento que você já foi nesse, nesse mundo do algoritmo, né?

Vinicius: Tá. Bom, contextualizando, assim, pegando desde o início da minha carreira, eu sou formado em Engenharia de Computação, então, eu tenho uma formação muito forte na área de tecnologia, né, então, eu trabalho tanto com a parte de algoritmos quanto na parte de hardware, né, então, eu faço as duas coisas, eu tenho a expertise de trabalhar nessas duas áreas.

Posteriormente, eu ingressei numa empresa de aviação, aqui do Brasil, onde eu trabalhei em Digital Engineering, onde lá eu cuidava de, de, do fluxograma da informação, então, todos os tipos de desenhos técnicos, de dados que a empresa gerava, que os colaboradores geravam, eu e a equipe, claro, não só eu, mas a equipe como um todo ali, cuidava de criar essa, esse meio para que os dados fossem possíveis de tramitar e atingir as pessoas necessárias e cumprir as regras, né, as regras de negócio e assim por diante. Aí, agora, eu, depois eu vim para Deloitte, né, aqui, e eu também atuei inicialmente em RPA, que é Robotização de Processos, Automatização de Processos via Robô, então, eu utilizava plataformas ou técnicas de low-code ou no-code, junto com código, lógico, para, para desenvolver soluções para diversas empresas, principalmente em questão de logística e tudo mais. Então, a gente entrava, fazia todo o processo de robotização e automatização, sem nenhuma necessidade de seres humanos, no caso.

E depois, eu, agora, acho que vai fazer uns dois anos e pouco, que eu atuo fortemente em Inteligência Artificial e Visão Computacional. Então, eu crio algoritmos de inteligência artificial para análise de imagens. Então, eu pego essas imagens, treino redes, né, de inteligência artificial para que encontre padrões dentro de imagens, de sinais, então, sinal de satélite, sinal de sonar e, assim, eu consigo gerar uma imagem através desse sinal e, aí, eu consigo extrair informações dessa, desse tipo de meio, né?

Entrevistadora: Ah, muito legal, é bem interessante.

Deixa eu te mostrar, isso vai ser bom, né, ter essa contextualização para entender, assim, até para eu entender os caminhos que você vai tomar na

solução, tá? Eu vou te passar um desafio, vamos dizer assim, e vou te contar o cenário fictício aqui, né? Imagina que uma empresa de tecnologia está com uma vaga aberta de, né, para um profissional pleno, é uma empresa de porte médio e ela contrata uma empresa que você está trabalhando, tá, no fictício, para fazer a divulgação desse anúncio em alguma rede específica, a gente não precisa se ater a qual rede, se é LinkedIn, mas imagina assim, sabe, Facebook, Instagram, quando aparece um anúncio para você, esse anúncio já foi direcionado para o seu perfil especificamente, né, assim, não é uma coisa padronizada para todos os usuários, então, o seu perfil já foi lido, tem suas preferências e tal, então, a ideia é que esse anúncio da oportunidade chegue a possíveis candidatos que vão dar fit com essa oportunidade, tá? Então, por exemplo, eu estou navegando em uma plataforma, sou de tecnologia, estou na Deloitte, trabalhando com tecnologia, enfim, aí esse anúncio vai aparecer para mim porque eles viram ali para alguma interação que eu fiz ou algum dado que eu forneci, que eu tenho a possibilidade de ocupar essa vaga. Então, a empresa que você trabalha, ela vai ser responsável por construir esse algoritmo para ler essas pessoas, tá?

Vinicius: Entendi.

Entrevistadora: Então, os requisitos que a empresa passa são bem, assim, genéricos, né, é uma vaga de tecnologia para profissional pleno com skills de desenvolvedora ou desenvolvedora full stack, com a formação em engenharia de software, tecnologia da informação, sistemas e áreas de ciências, né, geral, exatas, experiência de cinco anos e regime remoto.

Sim. E aí, assim, só isso que eles passam. Então, aí a nossa primeira etapa seria, como seria o planejamento e a concepção dessa solução? Em relação à definição de requisitos, como você define e refina esses requisitos que já foram passados, né, e como que você define para o algoritmo, né, esses requisitos, e quais dados você utilizaria, que tipo de dado você acha que é importante para responder aqueles requisitos que a empresa passou.

Eu vou voltar ali nesse slide, que aí você fica vendo quais são os requisitos.

Vinicius: Tá. Então, beleza, desenvolvedora full stack, informação, tecnologia, cinco anos de experiência, regime remoto.

Tecnologia para profissional pleno. Qual que é a outra slide, fazendo favor, só para eu ver os objetivos? Como você definiria os requisitos para o algoritmo? Que tipo de dados vocês utilizaram para alimentar o algoritmo? Tá. Você quer, esses requisitos, eles são mais técnicos ou mais subjetivos?

Entrevistadora: Podem ser subjetivos para você, você me explicando, assim, como que através dos dados, né, e dos requisitos, você vai direcionar esse perfil.

Vinicius: Escolher para as pessoas. Ah, tá, ok. Não, tranquilo, porque eu

já ia começar, não, tal linguagem, não sei o quê.

Entrevistadora: Mas pode falar também, fica à vontade. Vê, assim, o que for melhor para você construir o pensamento.

Vinicius: Bom, pode voltar para lá.

Entrevistadora: Vou voltar ao slide rapidinho.

Vinicius: Pleno, full stack. Eu desenvolveria um algoritmo, né, utilizando linguagem, provavelmente JavaScript, porque você consegue atuar tanto no front-end quanto no back-end, então você consegue ter as duas expertise ali.

O perfil das pessoas. Eu colocaria pessoas entre 25 anos mais, então eu teria um, não pegaria pessoas com, tá, colocaria 23, 24 anos mais, não pegaria pessoas muito novas, por conta desses 5 anos de experiência. Então eu, na relação de idade, caso eu tivesse acesso às informações de usuários, né, claro, teria ali a questão de você calibrar ali o algoritmo, colocaria 25 anos mais, vamos por aí, porque não daria tempo suficiente da pessoa ter 5 anos de experiência na área de tecnologia, né.

Desenvolvedor full stack. Eu buscaria pessoas que possuem algum tipo de postagem ou que buscam sempre nas plataformas visualizar essas questões de design juntamente com as questões de desenvolvimento, né, desenvolvimento back-end, né. Porque para ser um desenvolvedor full stack você precisa ter a expertise do front e do back, que assim, é meio complicado, porque é algo que, é um consenso normalmente entre o pessoal que é desenvolvedor, que é muito difícil você ter um cara bom nos dois.

É muito difícil. O cara teria que ter muitos anos de experiência e ter trabalhado muitos anos nessas duas áreas. Então eu colocaria, tipo, uma pessoa que busca sempre essa questão de user interface, user experience, juntamente com um cara que pesquisa bastante questões de desenvolvimento, de algoritmos em geral, né, de software.

Então uma pessoa que consiga trabalhar nessa, pesquise bastante essas informações. Sobre gênero, não colocaria nenhum marcador específico para gênero, pode ser pessoa de qualquer gênero, porque é independente. O que mais? É que aí eu não sei se eu teria acesso a, é só o acesso a informações de redes sociais.

Entrevistadora: É, assim, você pode, pode pensar...

Vinicius: É que a empresa vai criar um anúncio, né, para a pessoa se cadastrar, né, ok. Então aí eu colocaria, tal, a questão do... Eu tô pensando no anúncio, não tô pensando no algoritmo. Mas dos dados, né, esses são alguns da idade, os interesses da pessoa.

Então, por exemplo, que eu já tinha comentado, dessa questão dessas pesquisas, né. Gênero, qualquer gênero. Ah, a questão de regime remoto.

Buscaria também pessoas que possuem no perfil características em que ela possui alguma predileção para viagem ou para algo do gênero. Tipo, ah, que não tenha postagens relacionadas a muitos locais predeterminados, assim. Então, tipo assim, a pessoa só mora naquela região.

Então, muitas vezes, ela pode ter uma... Não digo uma necessidade, mas uma predileção a querer ficar naquele mesmo local. Entendeu? Então, uma pessoa que eu... Um critério que eu colocaria são postagens de viagem, de pesquisas, né, em questão de viagem. Porque tem muitos desenvolvedores que fazem aquela questão de nomad digital.

E o regime remoto ajudaria muito e facilitaria muito para essas pessoas, né. Então, colocaria esses tipos de dados, buscaria esse tipo de informação para que eu possa buscar esse profissional.

Entrevistadora: Ah, legal.

Vinicius: Tá. Acho que é isso.

Entrevistadora: E aí, um pouco mais assim, você já falou, né, mas se puder aprofundar nessa questão da escolha do método, tá.

Se você vai considerar aí técnicas de machine learning ou alguma coisa diferente. E se essa escolha de métodos está associada a algum desafio técnico que você imagina que pode enfrentar.

Vinicius: Bom, escolha de método.

Por se tratar de dados não... vai ter dado estruturado e não estruturado, né. Eu digo que vão ter dados tanto imagens quanto também pesquisas, né. Então, teria que ter... poderia ser um machine learning, né, para obter essas informações.

Então, não obter, mas para treinar a rede, né. Então, eu faria o input dessas informações numa técnica de machine learning. Agora, o modelo não poderia ser nada muito pesado, pelo fato de você estar querendo que ele... É, se bem que vai ter o servidor.

É difícil que a cabeça vai longe. Mas, colocaria alguma rede de deep learning. Supondo que eu tenha bastante poder computacional, colocaria alguma rede de deep learning.

Pensando que um SSD, que é um single-shot dimension, que é uma rede de machine learning. Que ela só passa uma vez pela rede e consegue extrair bem as informações ali. Os principais desafios técnicos no desenvolvimento do algoritmo seria a obtenção dessas informações.

Principalmente esse eu acho que é o maior desafio técnico. Porque você precisa de um grande número de dados para você criar um padrão. E aí você precisaria ter muitas pessoas com essa mesma característica que a empresa está buscando, né.

Então, ou você cria esses dados de forma artificial. Essas informações de forma artificial para você treinar a rede. Ou você então busca essas informações em algum banco de dados, em algum banco de dados mesmo.

Em algum banco de informação que você consiga integrar. Eu acho que esse é o maior desafio. Porque a implementação mesmo não é... Com o pessoal que está já atuando na área não é tão complexo assim.

Se tratando de dados que vão virar dados estruturados. São dados não estruturados, mas que viram dados estruturados depois. Como que o algoritmo foi integrado? Bom, todo o modelo você gera um treinamento, né.

Então você faz um input. Então você faz um pré-processamento que é o trabalho de você organizar essas informações. Então você extraiu essas informações, organizou elas.

Ou utilizando pandas, que é uma biblioteca de tabulação, né. Ou utilizando de forma manual. Aí depende de como que está a dinâmica, né.

A disponibilidade. Então você faz a estruturação desses dados. A partir disso você rotula eles.

Então você cria os rótulos em cada um dos pontos que você quer buscar. Da intenção. A partir disso você joga na rede.

Você passa ele ali no Deep Learning. Aí é interessante uma coisa é buscar alguma rede pré-treinada. Para que diminua o tempo de treinamento da rede.

Então uma rede que já tenha passado dados que você quer. Que você quer utilizar. Só que ao mesmo tempo tem que ser uma rede que, por exemplo, tem que ser dados muito similares.

Não pode ser dados que, por exemplo, vão dificultar o seu objetivo. Então, por exemplo, eu não posso pegar uma rede que vai identificar ou vai buscar um... Por exemplo, um de dados estruturados. Ah, uma rede que trabalha buscando idades.

Eu não posso colocar uma rede que busca idade e gênero. Porque isso aí vai acabar trazendo algum problema no resultado final. Então você tem que sempre pegar alguma rede que tenha muito parecido as informações.

Principalmente em dados tabulares. Em dados agora de imagem não há tanta necessidade. Pode ser, por exemplo, eu quero identificar carro.

Pode ser qualquer tipo de carro. Só se você quiser um tipo muito específico de carro. Aí você tem que fazer o ajuste fino.

Mas o ajuste fino você faz com esses dados que você está introduzindo. Então é só para ele pegar a ideia do que você quer. E aí você adiciona as mais informações.

E aí o pós-processamento, que é você gerar um modelo. Você já tem um modelo. Você fez o pré-processamento, gerou um modelo.

Ele passou, ele deu resultado. A partir dali você pode fazer uma integração via API. Onde você despeja essas informações que foram observadas.

Diretamente no algoritmo da web. Para fazer essa busca. Então eu realizei essa análise dessas informações.

Em todos esses tipos de postagem de várias pessoas. E aí eu crio um tipo de marketing. Ou um tipo de panfleto.

Ou algum tipo de post. Que obtenha todas essas características. E chegue nas pessoas necessárias ali. Entendeu? Mais ou menos isso.

Entrevistadora: Super bem explicado. E essa parte de testes. E a validação. Quais testes você realizaria. Se você consegue pensar. Alguma métrica de sucesso. Lembrando que você foi contratado para fazer. E é um produto que você está vendendo. E tem como mensurar esse desempenho. E se você considera feedbacks para as iterações do algoritmo. Feedbacks humanos, não humanos. Quais seriam?

Vinicius: Para a realização de testes. Eu acho que já poderia ser colocado diretamente. Inicialmente você poderia criar um ambiente de teste. Onde você gera esse tipo de informação. Baseado nos dados que você coletou. Você teve o output. E a partir dali você cria um teste interno. Então dentro da própria equipe. Dentro da própria empresa.

Você poderia criar um teste. Para ver se realmente está refletindo isso. Com uma medida de segurança. Para que não comece a disparar anúncios. Para pessoas que não tem nada a ver com esse tipo de informação. Então eu acho que seria interessante.

Um teste inicial. Onde você faria um teste com a própria equipe. Seria um pouco enviesado esse teste.

Porque você estaria fazendo com pessoas que já entendem. Como funciona o fluxo. Entretanto isso facilita para a correção de erros.

Porque pelo fato dela entender como o fluxo está fluindo. Como são as saídas. Você consegue identificar de uma forma mais fácil.

Erros esdrúxulos. Problemas muito latentes. É um erro muito crasso.

Então volta e recomeça. Posteriormente eu passaria para um teste já em bloco. Onde eu consigo disparar essa publicação.

A partir desses dados que eu fiz. Desse treinamento. Para um número X de participantes.

Para um número X de pessoas. Tendo em vista que eu tenho as métricas de publicação. Eu consigo tanto colocar feedbacks.

Nessa publicação que eu estou gerando. Então eu consigo extrair. Por exemplo.

Qual foi a pessoa que foi atingida. Se for um vídeo. Quanto tempo teve de retenção naquele vídeo.

Quantas pessoas clicaram naquele vídeo. Então eu consigo a partir disso medir. Quanto que foi obtido de informação.

De feedback tanto positivo quanto negativo. Para aquele tipo de interação. Então a partir disso eu consigo já ter os feedbacks.

De clicou no coraçãozinho. Ou clicou na publicação para abrir a página. Eu não sei.

Tipo. Uma página do emprego. A pessoa se interessou.

Ela vai até a página. Ela clica na página. Tudo isso você consegue medir.

Porque você consegue criar um dispositivo. Onde aquela publicação. Ela tem um link específico.

Que somente quem clicou naquela publicação. Vai para aquele link. Mesmo que seja uma homepage.

Da mesma empresa. Você consegue através disso detectar. Tipo.

O cara clicou aqui. Então eu vou saber que o cara veio por aquele e-mail. Então eu consigo ter esses feedbacks de interação.

E as métricas seriam. Retenção. Então quanto tempo a pessoa.

Depende se é um vídeo ou se é um post. Se é um post. Os likes já dão a retenção.

Os likes e os cliques. O vídeo você consegue. Tanto pelos likes quanto pelos vídeos.

Quanto tempo assistido. Porque a plataforma normalmente te passa esse tipo de métrica. Quanto tempo a pessoa se manteve ali.

Observando ali. Quanto tempo aquele aparelho ou aquele IP ficou ali. Naquele post. Seriam essas.

Entrevistadora: Tá. E você considera. Que são coisas mais quantitativas talvez. Que você trouxe. Você considera feedbacks da empresa. Pessoal do negócio em si. À medida que isso vai sendo testado. Rodado e tal.

Vinicius: Como assim? Feedback da empresa. Desculpa.

Entrevistadora: Da empresa que te contratou.

Vinicius: Então. Ah sim. Para eu desenvolver.

Entrevistadora: Os especialistas. Vamos dizer que tenha. Sei lá. Um especialista do RH. Que você tenha essa interface com eles. Do tipo. Para analisar os resultados. Sim. Sim.

Vinicius: Dá para. Colocar isso também. Isso também é um tipo de feedback. De interação do interno. Eu só pensei no externo. Agora no interno realmente. É quantas pessoas. A partir daí. Vai ter tipo. Quantas pessoas foram para a página. De recrutamento. E realmente realizaram o envio. De currículo.

Por exemplo. Comentaram na publicação. Isso daí é algo. Que o RH. Muito provavelmente vai. Utilizar para. Poder engajar. Numa possível contratação. Então eu acho que. Os Headhunters. Estariam observando. Essas informações. E olhariam o perfil. Dessa pessoa. Porque é claro. Vai ter. Os erros. Porque é uma rede. Ela nunca é perfeita. Ela nunca chega a 100

Tem que ter uma. Uma análise humana. Logo em seguida. Porque não dá. Para. Confiar totalmente. Numa análise. Só. Só máquina. Só. Algoritmo. Então tem que ter. Uma retroalimentação ali. Sim.

Entrevistadora: E você considera. Por exemplo. Quando você está desenvolvendo. Antes de gerar um modelo. Você considera. Trocar com os seus pares. Com pessoas que. Tenham. Assim. Nessa equipe. Trocar as decisões. Você gosta. Quando isso acontece. Ou você busca. Ou você. É mais assim. Isso. No caminho solo mesmo. Tomando decisões.

Vinicius: Não. Eu sempre. Eu sempre busco. Conversar com a equipe. É o mais indicado. Porque. Muitas vezes. Eu não. Tenho a visibilidade. Total. Do. Do que está sendo. Desenvolvido. Né. Às vezes. Eu tenho uma informação. Que pode ter. Que um colega meu. Tem uma outra informação. Que ele tenha conversado com o cliente. Ou que. Eu não posso. Eu não tenha. Ou não tenha escutado. Ou possa ter ocorrido. Algum problema. Uma falha de comunicação.

Então. É sempre importante. No. No. Principalmente. No pré-processamento. Porque no treinamento. Ok. O treinamento. Já está tudo certo. As informações. Já foram coletadas. Principalmente. No momento. Do pré-processamento. É sempre. Muito importante. A equipe. Está sempre. Se conversando. Falando. É. Delimitar. O. Escopo. Então. Ah. Sei lá. Identificação. De. De celular. O celular. Ele vai ter que ser. Dessa cor. Ele vai ter que ter. Tal tamanho. Ele vai ter. Que tal. Tal. Aparência. Então. Eu tenho que. Está. Conversado com a equipe. Dessa. Sobre essas características. Da mesma forma. Como um dado tabular. Eu tenho que. Ter tais características. Poxa. Tem tal característica. E tal. E característica x. E y mais. Esse y mais. Não estava previsto. Logo. Tem que ser conversado. Entre todo mundo.

Para ver. Ó. Isso aqui. Vai poder entrar. Isso aqui. Não vai poder entrar. Isso aqui. Pode. Não pode. Então. É muito importante. Para a gente. Ter um. Resultado. Fidedigno. Para. Porque. Senão. Você acaba gerando. Um monte de resultado. Que. Que. Que não vai. Resultar em nada. É. Vai ser um resultado ruim. É. O mais importante. De tudo. De tudo mesmo. É. O core. De toda. O processamento. É. É. O início. Dos dados. Se você tiver dado ruim. Vai sair dado ruim. Se você tiver dado bom. Vai sair dado bom. Sim. É principalmente. Essa dinâmica. E sempre. Para baixo. Se você tiver um dado ruim. Vai sair um dado ruim. Pior do que você introduziu.

Entrevistadora: Entendi.

Vinicius: Então. Então. É sempre essa dinâmica.

Entrevistadora: Entendi. E aí. Assim. Sobre a efetividade. O impacto. Na contratação. É uma coisa mais. Assim. Que. A sua opinião. Você considera que. Que é positivo. Usar esse tipo de algoritmo. Para fazer essa seleção. Esse recorte. É efetivo. Para já. Direcionar as pessoas. E. Se você. Tem alguma. Mensura. Algum impacto. Na contratação. Assim. Que que isso pode trazer. De diferente. Para esse processo. E qual é o impacto. Que pode gerar. Nessa parte da contratação. Vinicius: Bom. Eu acho que. Que a efetividade. Eu acho que é maior. Porque. Do que somente o ser humano. Por quê? Porque. Você vai fazer uma análise. Máquina ou humano. Você não vai ter só uma análise. Humano. Você vai atingir um número. Muito maior de pessoas. Então. Você vai ter um público. Muito maior. Porque a máquina vai conseguir. Disparar. E. Ter uma efetividade. De análise. Muito maior. Do que só. Uma pessoa. Realizando essa análise. Ou procurando efetivamente. Essa pessoa. Ou essa pessoa. Criar o post. E de forma orgânica. Digamos assim. Então. É um pouco enviesar. Aquilo. Mas é um viés. Bom. Porque é um viés. Que você. Tenha o objetivo. De atingir aquilo. Entendeu? É um viés. Objetificado. Ah. Eu tenho esse. Critério. Eu tenho que. Balizar. A minha solução. Para esse.

Para esse objetivo. Eu não posso deixar. É. Com arestas. Porque se não. Eu atinjo pessoas. Que não. Não é do interesse da empresa. E. E vai ser só. Um desperdício. Sabe. De recurso. Um desperdício de tempo. Porque não vai. Estar bem calibrado. E. Eu acho que é. Bastante efetivo. Porque a partir do momento. Que ela. As pessoas. Foram lá. E. E clicaram. Na publicação. Para a vaga. Aí entra o trabalho. Do Headhunter. Para. Observar. Se realmente. Aquilo está. Conducente. Se a pessoa possui. As. Expertises necessárias. Para aquele tipo de. De vaga. Que. Que. Que vai estar. Disponível. Né. Então. Eu acho que. Que é bem proveitoso. Utilizar isso.

Mas. Tem que ser. Junto com. Uma equipe. Humana. É. No. No pós. É. No pós acesso. Digamos assim. O pós trabalho. Da máquina. Para validar. Tudo aquilo. Que a máquina. Fez. Ou se. Está mal calibrado. É importante também. Que essa equipe. Conseguir dar feedback. Poxa. A máquina. Não. Não está me. Dando. As pessoas. Que eu. Que eu gostaria. Que que a gente pode. É uma retroalimentação. Que que a gente pode. Melhorar no algoritmo. Para ele. Conseguir atingir. Aquelas pessoas. Entendeu.

Entrevistadora: Entendi. Muito bom. Então. E deixa eu te perguntar. A gente falou de feedback. De troca. Entre. Né. Assim. Nas decisões. Sendo trocadas. Alinhadas. Em equipe. Você considera. Que como a gente está falando aqui. De um algoritmo. Que vai fazer um impacto. Na sociedade. Né. Assim.

Né. Vai escolher. Vai recortar. Um público. Que vai receber. Uma oportunidade de emprego. Né. É. Receber. O anúncio. Dessa oportunidade. Você considera. Que tem. Questões mais sensíveis. Né. Você falou. Durante a sua. A sua solução. Você falou de questões. Aí. Mais sensíveis. Mas você acha. Que seria interessante. Ter um tipo. Um checklist. Para você. É. Toda vez que você for fazer. Sei lá. Desenvolver um algoritmo. Que vai ter um impacto. Ou social. Uma coisa mais. Com uma camada mais sensível. Ali. De pessoas. Assim. Diretamente.

Em relação. A vida. Das pessoas. Não sei. Tem. Uma gama gigantesca. Né. De possibilidades. Você considera. Que seria importante. Ter um checklist. Para você. Assim. Não. É. Tentar. Não criar viés. Mas isso aí. É sempre. O seu. Seu. Sua intenção. Né. Claro. Mas. Muito das decisões. Podem partir. Das suas experiências próprias. Né. Ainda mais assim. Contratação. Você fica pensando. Né. Quem eu gostaria de contratar.

Ou como já me contrataram. Né. Uma coisa mais pessoal. Você acha. Que seria interessante. Ter um checklist. Para balizar. Questões. Importantes. Para você pensar. E. Tentar blindar. Essa questão. De. De viés.

Vinicius: Olha. A resposta. É. Depende. Depende. Bastante. Porque. É. Se os critérios. Forem. Muito bem. Definidos. Eu acho. Que ok. Você pode. Por um. Checklist. É. Assim. Eu acho importante. Ter um checklist. Genérico. De coisas. Que realmente. Por exemplo. A raça. Cor. Essas coisas. Eu acho que. Realmente. Tem que ter. Para ter um. Uma equidade. No processo. Né. Então. Tipo. Ah. É. Não só. Para que o algoritmo. Não. Não envia apenas a publicação. Só para pessoas brancas. Ou só para pessoas indígenas. Ou só para pessoas negras. E assim por diante. Então. Você tem que ter um checklist. É. Quantidade. De. De. De. Entradas. Né. Porque tem que ser o. Igual. Igual. Igualitário. Ali. E. De gênero. Também. Para que. Não. Gere nenhum. Desbalanceamento.

Então. Os dados. Tem que estar. Balanceados. Para. Diversos grupos. É. Da sociedade. Agora. Se. Algo. Muito específico. Ou que. Por exemplo. É. É que. Nesse caso. Específico. Que você me passou. Eu acho que. O checklist. Genérico. Já. Solucionaria. É que. Eu estava falando. Em casos específicos. É que. Às vezes. Tem vieses. Que são naturais. Que são. Coisas. Vindas. Do tipo. Teve uma pesquisa. A pesquisa. Constatou isso. E é isso. Tem que aparecer. Esse. Viés.

Porque. Algo. Natural. Que existiu. Agora. Para. Ser. Algo. Natural. É algo. Que você quer combater. Ou é algo. Que você quer. Que a empresa. No caso. É. Mude. O perfil. E tudo mais. Aí sim. Vamos. Deixar tudo igualitário. Eu acho importante. Sim. Ter esse checklist. Mas. De uma forma. Mais

genérica. Sim. De questões. Mais sensíveis.

Entrevistadora: Mais direcionador. Mesmo.

Vinicius: Isso. Agora. Para. Coisas específicas. Onde tem um viés. Que. Tranquilo. É. Você manter esse viés. Porque é algo natural. Que vai existir. Não. Não é algo. Que vai. Trazer problema. Para alguém.

Entrevistadora: Tá legal. Foi isso. Deixa eu parar de gravar aqui. Muito obrigada. É. Eu.

B.11

Alice

Entrevistadora: Ou a transcrição.

Alice: Você precisa que a câmera esteja ligada?

Entrevistadora: É uma boa, se você... Tá, deixa eu começar aqui. Primeiro, você tá vendo a tela já?

Alice: Tô vendo.

Entrevistadora: Eu vou pedir só para você fazer uma contextualização contando a sua experiência, principalmente com algoritmos ou nessa área de análise de dados, enfim, e um pouco do que você faz hoje. Não precisa entrar a aprofundar em clientes, né, no projeto, mas só uma geral mesmo.

Alice: Beleza.

Então, Larissa, eu sou geofísica de formação, né? O meu primeiro contato com o desenvolvimento de algoritmos foi na faculdade, no terceiro período, com lógica de programação, e aí eu, na época, eu peguei um trauma profundo de tecnologia e programação. Achei que não era para mim. Segui minha área, fiz mestrado na minha área, não me envolvendo com essa parte de desenvolvimento de códigos, né? Depois que eu finalizei o mestrado, eu comecei a mudar de opinião em relação a isso.

Pensei que eu poderia sim aprender essa parte de algoritmos. Foi quando eu fiz dois cursos. Um foi um curso mais simplificado, de ações afirmativas para mulheres.

Eu tive uma visão mais generalizada sobre front-end, HTML, CSS. Então, foi meu primeiro contato, assim, depois do terceiro período da faculdade. Acho que foi, o quê, uns sete anos depois, bastante tempo.

Enfim, e aí depois eu fiz um outro curso de três meses, mas voltado para a nuvem, e depois entrei na Deloitte, né? Foi quando eu comecei a ter um contato maior com a parte do desenvolvimento de código, mas voltado para análise de dados e ciência de dados, né? Então, atualmente, meu papel na empresa com algoritmo é manutenção de soluções que já existem, focando em melhorias e inovação, né? Então, a princípio, é isso. A ideia é que futuramente eu consiga

atuar na parte de viagem negativa, mas, a princípio, é pelas demandas do projeto, né? O meu foco é mais manutenção.

Entrevistadora: Maravilha, legal.

Opa, eu vou te passar um cenário fictício, e aí você me fala qual solução você teria, tá?

Alice: Tá bom.

Entrevistadora: Imagina que uma empresa de tecnologia, de porte médio, tá com uma vaga aberta para anunciar, e eles pretendem contratar um profissional de tecnologia, né? Então, o que eles fazem? Eles vão até uma empresa, que é a que você trabalha, tá? Nesse cenário, para fazer um algoritmo para ajudar a divulgar essa vaga específica, para aqueles perfis específicos que eles acham que vão, que vai ter match, tá? Então, assim, eles passam esses requisitos. Eu vou falar quais são, quais são os requisitos, de uma maneira muito abrangente, né? Alto nível aqui, e dizem para você, só anuncia, faz um anúncio e só anuncia para quem você acha que tem perfil, tá? Então, aqueles anúncios pagos que já tem um recorte de público, né? E aí, a sua função é fazer esse algoritmo que vai selecionar essas pessoas.

E aí, quais são os requisitos que eles passam? É uma vaga de tecnologia para um profissional pleno, com skills para desenvolvedora ou desenvolvedora full stack, formação, engenharia de software, tecnologia da informação, sistemas e gráficos, uma experiência básica de cinco anos, né? Mas, assim, isso tudo aqui você vai ter que, vamos dizer, entender melhor, né? E sendo regime remoto, tá? Então, eles passaram esse cenário e aí a primeira fase seria você fazer o planejamento para a concepção desse algoritmo. Então, como você faria a definição, o refinamento desses requisitos, especificamente, que eu falei? E quais seriam os dados utilizados? Que tipo de dado você utilizaria e que tipo de informação você precisa para responder aos seus requisitos?

Alice: Tá, então, deixa eu pensar aqui.

Entrevistadora: Quer que eu volte lá para você ver os requisitos?

Alice: Por favor.

Bom, a princípio, né, eu colocaria dados de formação dos candidatos, porque assim daria para fazer o refino do tempo de atuação, da formação que essas pessoas têm, né? E das skills, porque com o currículo a gente consegue ter tudo isso. Então, no algoritmo, eu acho que eu conseguiria colocar um filtro mais específico, mais uma margem, acredito que tem que ter essa margem, né? Para que tenha esses requisitos, principalmente em relação ao tempo exigido de experiência e a formação. Em relação às skills, palavras chaves que teriam que ter presente.

Então, nesse algoritmo, seria importante que a gente conseguisse ter um

match, né, de algumas palavras específicas. Por exemplo, no caso, um full stack. Isso precisaria estar no currículo.

Então, a princípio, acho que o principal dado que eu teria seria o currículo dos candidatos.

Entrevistadora: E aí, você buscaria, tipo assim, não tem problema, você não precisa fechar em plataforma, tá? Do tipo, ah, precisa ser LinkedIn.

Alice: Não, dá para fazer uma busca ativa.

Isso. Entendi.

Entrevistadora: Mas assim, não precisa se restringir pelas questões burocráticas, tá? Imagina que tudo está liberado e tal.

Alice: Eu começaria, se eu sou uma pessoa que estou recrutando, né, eu tenho que pensar apenas no algoritmo ou não?

Entrevistadora: Não, pode pensar... Ah, tá. No todo.

Alice: Tá bom.

Então, eu começaria recrutando através dos grupos que eu tenho, que fazem parte desse contexto de tecnologia. Então, eu começaria divulgando a vaga a partir daí, definindo os requisitos, especialmente a formação e experiência. E isso no grupo de WhatsApp, Telegram, né, que são mais voltados para a área de tecnologia, especialmente para as vagas afirmativas.

E aí, depois, uma publicação, né, no LinkedIn, para poder expandir mais, caso não tenha tido um número legal de inscrições da outra forma mais direta, né, que seriam através dos grupos.

Entrevistadora: E nessa divulgação no LinkedIn, por exemplo, se você fosse colocar, só mostra para tais pessoas, quais dados que você achava que seria importante, assim?

Alice: Bom, pessoas que, a princípio, isso daí é um pouco pessoal, sabe? Pessoas que estariam com Open to Work, para mim, esse é um requisito bom para prioridade ser as pessoas que, no momento, estão desempregadas. E as pessoas que estão buscando algum tipo de recolocação no mercado, né? Não teria mais esse conforto.

Entrevistadora: E, peraí, deixa eu voltar para cá. Sobre o desenvolvimento em si do algoritmo, né? Que métodos você utilizaria? Seria Machine Learning, outro método qualquer, né, que você acha que vale para essa automatização, né? Lembrando que, assim, esse anúncio vai aparecer para as pessoas que você acha que tem o perfil. Já fez a leitura inicial, né? E aí vai fazer esse anúncio ativo.

E aí, os desafios técnicos, se esses desafios que você consegue mensurar podem, inclusive, direcionar a escolha do seu método. E essa parte de integração e implementação, pensando que fosse um algoritmo que você vai colocar

em uma plataforma. Mas isso é o menos importante, assim, é mais a escolha do método e dos desafios.

Alice: Então, eu acho importante que tenha essa questão do Machine Learning, né? Porque é algo que é para facilitar, né, esse controle e o recebimento. Só que eu não acho que deveria ser exclusivo a isso, porque, no fim das contas, a gente está lidando com pessoas, né? Então, eu vejo que, às vezes, dependendo da forma que esse algoritmo vai ser construído, isso pode ser mais excludente, né? Então, eu acho que eu uniria duas formas, uma mais tradicional, que é, literalmente, a gente recebe, coloca os requisitos, recebe a candidatura da pessoa e avalia, ou a pessoa nos processos do RH, né? E a outra parte seria com o Machine Learning, ver se tem método das palavras ou método dos requisitos, né? Então, assim, eu optaria por seguir essas duas frentes de maneira paralela, porque eu, na construção do algoritmo, né, a gente tem um processo que pode ser, não falho, mas pode ser realmente limitante. Às vezes, eu tenho uma pessoa muito boa e, na hora de escrever, por exemplo, ela não colocou a palavra chave X, só que ela foi uma pessoa que vai se adequar a esse perfil, né? Então, se eu sou uma pessoa do RH, que eu estou procurando uma pessoa, que ela não vai ser apenas uma caixinha, que eu também tenho que abrir os meus olhos para isso, né? Eu acho que eu já, não sei se eu já comentei com você, mas eu já fiz parte de um RH, mas foi na época da empresa Júnior, né? Eu lidava justamente com essa parte de aceitar as pessoas, trazer as pessoas para a empresa e tudo mais, né? E a gente tem formas muito automatizadas de lidar com isso, só que nem sempre isso vai ser suficiente, porque a gente está buscando mais a longo prazo.

Então, por isso que eu iria fazer duas frentes.

Entrevistadora: Legal. E em relação aos testes para realizar e validação, né, em métricas que você pode levantar, e lembrando que é um produto, né, para ser vendido.

E esses feedbacks, você considera feedbacks humanos, não humanos, para fazer a iteração do algoritmo?

Alice: Eu não sei se eu compreendi totalmente. Calma aí, eu estou tentando reler aqui. No caso, o feedback são das pessoas que se candidataram para as vagas?

Entrevistadora: Não, para você ir simbolizando o desenvolvimento desse algoritmo, né? Então, por exemplo, se você, a partir dos testes, está rodando, está selecionando pessoas, aí, por exemplo, um exemplo, tá? O RH pode dizer, não, os candidatos que vocês estão direcionando, não, tá bom, tal, é um feedback.

Ou vocês, através de estabelecimento de alguma métrica, você pode

conseguir capturar feedbacks automáticos, sabe? Que também pode resultar em iterações.

Alice: Eu acho que essa vai ser a etapa, digamos assim, crucial, né? Porque a gente vai conseguir calibrar exatamente para o que a gente procura, não apenas na parte automatizada, mas no que realmente a empresa está precisando ali, né, na hora da contratação. Então, por exemplo, se a gente está fazendo essa campanha de divulgação de vaga, umas métricas importantes vai ser, com certeza, quantas pessoas que se candidataram, quantas pessoas que se candidataram estão preenchendo os requisitos que a gente tinha definido antes, independente de ter sido pelo machine learning ou pela forma mais manual, digamos assim, dos processos, né? Então, no final das contas, tudo isso entraria ali nas métricas, né, para avaliar o sucesso dessa campanha.

Em relação ao feedback, eu acho que a quantidade, né, não necessariamente ela vai ser a qualidade, mas isso é muito individual do que o RH, em essência, está buscando. Então, para isso, precisaria ter toda a equipe, né, reunida para definir, afinal, o que vale realmente são esses requisitos? Porque se esses requisitos, eles devem ser considerados para que as pessoas participem do processo, mas ele não é suficiente, certo? Porque senão todo mundo que teria o requisito conseguiria entrar. Então, o que vai ser mais diferencial ali? A forma que a pessoa vai se comunicar, algum título que essa pessoa tenha, projetos que ela já participou, independente de ser empresa ou não.

Então, assim, esses outros critérios vão precisar entrar para que isso seja um método de classificação, né, depois que a gente passa por essa peneira inicial. Então, o algoritmo talvez deva, assim, ser modificado para que a gente vá acrescentando esses novos critérios, né, para conseguir ir deixando o processo um pouco mais robusto do que a gente quer no final.

Entrevistadora: Tá.

E aí, em relação a uma coisa mais no final, assim, né, sobre a efetividade, você acredita que esse esse tipo de solução pode ser efetiva, de fato, na hora da avaliação e recomendação desses candidatos? E se você acredita que esse tipo de processo tem algum impacto na contratação? Aí você pode considerar que, assim, são pessoas, são vidas, né, são questões de pessoas pertencendo a vagas, né, assim, mas questões sociais mesmo também você pode avaliar se tem esse impacto, né, na contratação.

Alice: Bom, então, em relação à efetividade, a gente tem que seguir de uma forma talvez mais objetiva, né, então primeiro seria definir o que seria um algoritmo estar efetivo. É ele conseguir apontar para todas as pessoas que estavam nos requisitos? Se for sem beleza, a gente vai conseguir ter, avaliar isso de forma mais direta.

Mas, assim, eu compordo que a gente tenha que ter uma visão social do que a gente está propondo, né, porque realmente estamos lidando com pessoas, com vidas, com sonhos, né, então eu não acho que dê para que o algoritmo, ele faça o papel e substituir um RH com um olhar mais humano. Porque, por exemplo, eu fico pensando, temos uma pessoa que, sei lá, estou a vida inteira, e aí na construção do correto dessa pessoa, talvez ela tenha perdido uma palavrinha-chave que ali tenha feito diferença na hora do algoritmo apontar essa pessoa aqui. E não necessariamente quer dizer que essa pessoa não está apta para o trabalho, né, então eu acho que a gente vai ter que ter um cuidado para que não tenha só esse viagem mais direta.

Ah, tem isso aqui, então, bom, perdi. Tendo esse cuidado, o impacto que isso pode gerar no processo de contratação da empresa, é, a princípio que eu vejo como algo muito positivo é a agilidade, né, no processo, que seria algo, assim, mais forte, mas os recursos humanos, eles estão ali para que a gente lide com as pessoas. A tecnologia e os algoritmos, eles deveriam agir apenas como uma forma de facilitador, né? É a visão que eu tenho, talvez seja um pouco conservadora, da minha parte, ainda que eu trabalhe na parte de algoritmo, né, porque é, no final das contas, estamos trabalhando com pessoas não cópias, né? Então, eu acredito que tem que ter um cuidado da empresa, já nesse processo, de como não tornar tudo muito burocrático, enferrujado, não pensando apenas na aceleração do processo, que eu acho que isso é bem-vindo para todo mundo, mas sem que isso seja o que vai dizer que aquela pessoa ali, que às vezes ela está, assim, apta para o trabalho, porque um algoritmo disse que não.

Então, eu acho que tem que ter um cuidado na hora de definir esses critérios e sempre ter uma margem para que a gente não seja excludente nas nossas escolhas, né?

Entrevistadora: Sim, muito bom. Legal, não? Muito bom. É sempre muita discussão, né, assim, que gera e porque está afetando questões de sociedade, né, e de acesso, né?

Alice: Uma pessoa vai mudar de vida se ela não tem nem a oportunidade.

Entrevistadora: Sim. E aí, deixa eu te perguntar, quando você está desenvolvendo os algoritmos e tal, tomando essas decisões, você costuma consultar seus pares, né, assim, para direcionar, você acha que aqui e tal. Você acha que seria legal, se tivesse um framework ou um checklist, vamos dizer assim, quando existem questões mais sensíveis, por exemplo, esse algoritmo claramente vai afetar vidas, vamos dizer assim, né? Então, são questões mais sensíveis.

Você considera que se tivesse um checklist, por exemplo, para você

ir pensando em situações, enquanto você está modelando, desenvolvendo o algoritmo, pensando em situações que podem, sei lá, trazer questões de diversidade, inclusão, afirmação, né, enfim, todas as coisas mais de pluralidade das relações, então. Você acha que seria legal e você, de fato, acha que seria efetivo, assim, você usaria?

Alice: Eu acho que é algo que nós temos obrigação de pensar, entendeu? Eu não acho nem que seja uma questão de legal, eu acho que é realmente uma obrigação, sabe, de ter essas considerações, porque senão nós, a resposta vai ser sempre a mesma, né, e não teria mudança. Então, a forma que eu vejo de um checklist seria uma forma de tentar minimizar esses efeitos ou nos obrigar a, na construção desse algoritmo, pensar em outras possibilidades, digamos assim.

Então, eu acho que seria essencial, assim, uma boa saída.

Entrevistadora: Legal. É isso, muito obrigada, viu? Você mandou muito bem, você estava toda aí, ah, não sei se eu vou, vou saber.

Alice: É difícil.

Entrevistadora: É, não, mas assim, você mandou super bem.

Alice: É, sua pesquisa é muito interessante.

Entrevistadora: Que bom, você acha. Sim. Obrigada.

É, eu acho que, assim, desde que eu comecei aí a entrevistar, né, as pessoas e tal, minha cabeça já mudou um pouco.

Alice: Sério?

Entrevistadora: É, isso é interessante. Eu estava querendo direcionar para uma análise, mas aí eu comecei, por isso que eu tô enfiando essas perguntas no final também.

Alice: Legal.

Entrevistadora: Inclusive, esse formulário que eu vou te enviar aqui, na hora que você puder responder, por favor.

Alice: Ah, já respondo.

Entrevistadora: Ah, sem pressa. Esse formulário eu tô enviando depois, porque aqui explica um pouco da minha pesquisa. E aí, algumas pessoas ficam enviesadas na hora de falar, entendeu?

Alice: Ah, por isso que você tá enviando depois.

Entrevistadora: Legal. Aí tô enviando depois, eu percebi. É legal, é muito legal, assim, eu tenho prazo insano, né? É isso que você tá pensando, né, Larissa? É até o meio do ano? Nada, é, na verdade é, é até junho, só que para eu apresentar em 20 de junho, eu tenho que estar com o texto pronto em 20 de maio.

Alice: Nossa, é um mês antes, né, gente? Então você tem, assim, uns 20 e poucos dias para escrever, né? Ah, dois meses, né?

Entrevistadora: É, assim, até o final de maio eu consigo, então vai ser abril e maio, e aí eu vou tirar férias daqui a duas semanas, e vou ficar duas semanas imersa. Só escrevendo. Só escrevendo, se eu pudesse desligar o celular, eu faria, assim, para não...

Alice: Não, legal, para ficar bem concentrada.

Entrevistadora: É porque esse negócio tá cortando...

Alice: Você tem facilidade para escrever?

Entrevistadora: Não, não, assim, na verdade eu não tenho dificuldade de escrever, eu tenho dificuldade de concentração. Ah, sim, eu te entendo. Foco.

Quando eu consigo focar, eu vou, sabe? O negócio é que, assim, esse negócio de trabalhar e chegar à noite e pegar para fazer... Eu não consigo mais, assim, porque também são outras responsabilidades que você tem no trabalho, né? É mais complexo. Então, eu não tô conseguindo. Infelizmente, eu vou ter que tirar minhas férias para trabalhar, mas vai acontecer.

Alice: Ah, mas pensa que depois de 20 de maio vai ser liberdade.

Entrevistadora: Com certeza. Obrigada, viu?

Alice: Nada que isso, eu que agradeço. Boa sorte.

Entrevistadora: Tchau, tchau.

B.12

Ana

Entrevistadora: Tati tá em português, tá? Também vou colocar do Not Disturbed aqui, que o negócio tá pipocando. Pronto, tá em português. Eu vou compartilhar com você a minha tela, que é assim, são só algumas algumas frases, né, para ir pontuando o corpo, né, da nossa entrevista, mas não tem nada específico assim.

Se você quiser compartilhar alguma coisa, compartilhar sua tela com alguma ferramenta que você queira mostrar, ou alguma plataforma, qualquer informação, fica à vontade, tá? Deixa eu abrir aqui, aí a gente começa. Você tá vendo minha tela já? Tá bom. Primeira coisa, eu queria até para gente entender melhor aqui o contexto, né? Eu queria que você, se puder, contar um pouco da sua experiência com algoritmos, experiência profissional ou acadêmica, e um pouco do que você vem desenvolvendo no seu trabalho atual.

Ana: Tá bom. Bom, a minha experiência com algoritmos, ela começa antes daqui da Deloitte, né, eu trabalhei por sete anos na IBM, e lá eu comecei a atuar como cientista de dados lá para 2019, não, 17, confundo os números ímpares, 2017, por aí eu comecei. Fiz um trabalho mais especificamente em

2019 nessa área com algoritmos, trabalhei na área de compliance e de análise de riscos, principalmente na área de compras, né? Então, dentro da área de compras de uma empresa, a gente vai ter lá as notas fiscais, né, dos clientes e tal, então eu trabalhei nessa parte para a gente tentar analisar o risco de ter algum, alguma fraude ou algum desvio de compliance nessas compras, né, no processo de compras dentro da empresa.

Então, foi nesse projeto que eu comecei a adentrar mais esse mundo dos algoritmos, e também foi quando eu entrei, assim, numa parte um pouco mais acadêmica, quando comecei a fazer uma pós-graduação em ciência de dados para poder também entender e lidar melhor com esses algoritmos, né, e a gente ia começar a aplicar ali. A princípio, a gente trabalhava só com regras de negócio mesmo, então tinha que obedecer uma certa regrinha para a gente impor ali se aquela, se uma nota fiscal tinha algum risco ou não, e depois de um tempo nessas regras, a gente decidiu aplicar alguns, fazer alguns estudos, né, utilizando inteligência artificial para ver, usando os algoritmos, né, de machine learning, para ver se a gente conseguiu um resultado melhor sobre quais notas fiscais que eram possíveis de ter um desvio de compliance ou um risco maior ali, né, associado. Então, acho que essa foi a minha base inicial, né, então foi dentro do mundo corporativo, e aí eu migrei ali para fazer uma pós-graduação para poder me localizar melhor, né, e saber exatamente com o que eu estava lidando, e então fiquei bastante tempo lidando com esse projeto, lidando com essa parte de fraude em nota fiscal.

Depois disso, não fiz nenhum trabalho muito relevante, assim, eu mudei de empresa, não fiz muito trabalho relevante na área com algoritmos, e aí eu entrei aqui na Deloitte e trabalhei já com a parte de visão computacional, e agora estou trabalhando na parte de engenharia de poços, né, é mais uma parte, assim, vários, vários, são vários projetos que a gente está pegando, a gente não está fazendo, a gente está pegando projetos que já foram desenvolvidos para a gente poder melhorar, produtizar, enfim, entregar o resultado final para o usuário, é nessa parte de o risco do poço, de ter que ter um retrabalho nesse poço, de petróleo, uma categorização também dos poços, como se fosse tipo uma clusterização desses poços, então diversos tipos de trabalho relacionados aos poços em si.

Entrevistadora: Ah, legal, muito bom. Então, quando você estava falando sobre essa questão de entender um padrão, né, nota fiscal, que pode dar mais chance de fraude ou não, né, é mais assim com classificação que você foi, que você trabalhou, algoritmo de classificação?

Ana: Isso, a gente trabalhou com algoritmo de classificação ali em cima das regras de negócio, então a gente tinha diversas regras de negócio, ela devia

dar umas 20 assim, e aí a gente analisava o resultado dessas regras, então se a regra falava essa nota fiscal pode ser problemática, então se a gente ia analisar ali a combinação dessas regras, e aí ia ter um risco associado em relação a cada regra que pegava, né, importante também frisar que a gente tinha esses dados, esses dados eles eram revisados por pessoas mesmo, por humanos, né, então todas as notas fiscais elas eram revisadas ali por humanos, e aí elas tinham uma aprovação, uma rejeição associada, a gente usou também esses dados para a gente poder criar um histórico ali e fazer um treinamento supervisionado.

Entrevistadora: Entendi, legal, muito bom. Tá bom, eu vou passar o contexto, aí não sei, se você de repente tiver alguma experiência que você já passou, que vale dizer que também que cabe para essa solução, fica à vontade também. Vamos imaginar um cenário hipotético, até só dizer um pouco do método que eu vou usar, a gente chama de *directed storytelling*, que é eu peço para você, eu passo para você um cenário, e aí você vai me dizendo realmente quais ações, o que você faria, pensando no passo a passo para resolver essa solução, e vai me dizendo, tá, então é uma coisa bem assim, não tem problema se você quiser mudar de ideia, ou se é falar alguma coisa que você não tem certeza, fica tranquilo, tá, é esse o exercício mesmo.

O cenário, então, hipotético que a gente criou é, imagina que uma empresa de tecnologia de porte médio contrata você, que está trabalhando numa empresa de divulgação de vagas, tá, então essa empresa de tecnologia, ela precisa divulgar um anúncio para atrair pessoas para se inscreverem numa vaga de emprego que está aberta, né, para essa empresa. Então, eles passam apenas alguns requisitos, que são esses aqui, né, que é uma vaga de tecnologia para um profissional pleno, algumas skills, fala de uma maneira, né, bem abrangente, que alcance um desenvolvedor ou desenvolvedora full stack, com formação em engenharia de software, tecnologia da informação, sistemas, né, mais na área de exatas, e experiência de cinco anos de experiência, mais ou menos, e o regime remoto, tá, então a empresa te manda esse contrato para fazer um anúncio, e como seria esse anúncio? Imagina assim, na época do Facebook, até Instagram mesmo, né, a gente tá navegando, aparecem anúncios que têm a ver com o nosso perfil, de alguma forma eles já foram direcionados por algum algoritmo, que entendem que a gente vai capturar aquela informação que está sendo veiculada, né, patrocinada ali. Então, a ideia é fazer uma busca, montar esse anúncio com um detalhamento específico de perfis que possam cumprir, né, ser cumpridos para ocupar essa vaga específica de trabalho, tá, aí assim, se vai ser no Facebook, no LinkedIn, ou qualquer, né, você não precisa se fixar na plataforma específica, né, porque é uma coisa mais hipotética mesmo, mas, de certa forma, esse anúncio só vai aparecer, você vai organizar, né, um algoritmo

que ele vai fazer uma, é, vai entender o perfil das pessoas, e esse anúncio só vai aparecer para aquelas pessoas que você acredita que podem ocupar essa vaga, você, né, representando essa empresa.

Daí, a primeira, depois, se quiser, eu volto aqui nessa slide, mas a primeira etapa é como você faria o planejamento e a concepção desse algoritmo. Primeiro, em relação à definição de requisitos, como você busca mais definição, como é que você pode fechar nesses requisitos ali, ou ter, de repente, mais elaborados, e esses dados que você vai utilizar, quais seriam, né, os dados que você vai utilizar para cumprir esses requisitos.

Ana: Nossa, eu não sei, eu nunca trabalhei com dados de rede social, assim, então eu não faço ideia como que é, para saber que dado que eu posso utilizar, ainda mais tratando de dado, é, de dado pessoal, né, então...

Entrevistadora: Perdão, é só para dizer assim, você não precisa se preocupar com as questões técnicas e éticas, não, tá? Vamos dizer que, no mundo ideal, LinkedIn, você pode usar qualquer dado que esteja ali, não tem problema em relação a essas restrições, né, específicas.

Ana: É, acredito que a rede mais fácil de se conseguir isso seria no LinkedIn mesmo, porque lá a gente tem os dados de trabalho, né, o dado do que a pessoa faz da vida, enfim, no Instagram e nas redes sociais, eu acho que estaria mais associado às páginas que essas pessoas seguem, então, ah, se é uma pessoa que trabalha com TI, então ela vai seguir páginas que são afins a essa área, né, então vai seguir ali alguma página de memes de área de TI também, né, mas acredito que no LinkedIn seja até mais fácil para conseguir até captar as pessoas que estão buscando, né, realocação no mercado de trabalho, que já sinaliza ali, né, mas, é, o que mais?

Entrevistadora: Quais dados? Eu vou até voltar nesse outro slide aqui, porque aí, opa, quais dados que você acha que você buscaria para ver se está convergindo com esses requisitos, né, assim, o stack, formação, quais dados?

Ana: Provavelmente, eu pegaria ali o título, né, do que a pessoa faz, para ver se está ali correlacionado, se é afim a formação, dá para fazer também uma leitura das experiências ali, né, no LinkedIn a gente sempre tem as experiências listadas ali, para ver quantos anos mais ou menos de experiência a pessoa tem, essa parte do regime remoto, eu acho que é de menos, porque, né, isso a gente, sendo remoto, a gente não precisa ficar se preocupando com que região que a pessoa está, não ser que tenha algum, algum problema, tipo, ah, de que tem que ser pessoa alocada no Brasil, tipo, não pode ser fora do país, né, nesses casos, mas, então, não precisa pegar uma localização ali, mas, no caso, pegaria, então, o título da pessoa, né, ali que aparece, né, sempre quando a gente procura no LinkedIn, abaixo, é, a formação também, que sempre tem,

é, e aí a experiência, né, então, tentar fazer uma leitura ali da experiência, é, quanto tempo de experiência que aquela pessoa tem em cargos correlacionados, provavelmente seria algo do gênero.

Entrevistadora: Tá legal, para ver se vai batendo com esse perfil específico aqui, né?

Ana: Isso, isso.

Entrevistadora: Tá legal, daí assim, em relação ao algoritmo mesmo, né, qual método você escolheria para fazer essa, essa busca, né, esse filtro dessas pessoas com os requisitos específicos, é, e possíveis desafios técnicos que você pode alcançar, né, então, assim, isso pode, de repente, motivar a sua escolha do método.

Ana: Nossa, é difícil, eu nunca trabalhei com nada parecido, então, eu não faço ideia de como que é fazer mapeamento de, de, de, de redes, assim, é, a coisa mais parecida que eu já fiz foi, é, NLP, que eu fiz um, é, eu coletei dados de, tipo, de, do Rotten Tomatoes para poder avaliar filmes, mas, mais parecido, assim, que eu fiz, que lidava com texto, foi isso, que, aí, a gente vai ter que pegar texto, né, então, sei lá, talvez, faria alguma coisa utilizando um NLP, ali, para a gente procurar por, por nomes parecidos, ou, então, até fazer um fuzzy match para ver, é, então, beleza, eu tô procurando desenvolvedor full stack, então, faz um NLP, ali, associado a, a um fuzzy match, para a gente poder encontrar palavras similares, né, termos similares, e, o que mais? Desafios técnicos, nossa, não faço ideia, provavelmente, também tem a questão da língua, né, a gente vai ter, ali, então, é, é, os termos em inglês, os termos em português, importante a gente, também, frisar, é, se o, que língua que o candidato precisa falar, se ele precisa falar português, se ele precisa falar português, então, aí, já é um desafio maior, porque, é, geralmente, as pessoas, pelo menos, as brasileiras, elas, eles estão lá no linkedin, os idiomas, e, tipo, geralmente, nunca tem português lá, porque elas já assumem que, tipo, um recrutador humano vai olhar o perfil dela e vai pensar, é, esse é brasileiro, vai, é, então, talvez, pegar essa parte da nacionalidade, também, para entender o idioma que vai ser necessário, e, também, na busca, né, então, tem o idioma do, que a pessoa precisa falar e tem o idioma da busca, né, que, então, tem gente que coloca lá, é, sei lá, é, full stack developer, tem gente que coloca desenvolvedor, full stack, então, a gente tem que trabalhar nesses termos, também, é, mas eu, eu, eu acho que NLP já faz isso, a gente tem como fazer essa tradução dos termos para poder, é, centralizar mais essa busca, que é mais, tem que tomar cuidado, também, para não deixar essa busca muito ampla, então, assim, pegar só desenvolvedor, pegar só desenvolvedor, eu não sei exatamente o que um full stack faz, mas pegar, sei lá, um desenvolvedor Cobol, sabe, não tem nada a ver, assim, tem

que ter esse cuidado, também, de não deixar essa busca muito aberta, o que mais? Ai, não consigo pensar em mais nada, assim, eu acho.

Entrevistadora: Não, já é bastante coisa que você disse, e aí, na parte de testes e validação, né, se você consegue pensar, assim, que tipo de teste você faria, é, para saber se o algoritmo está respondendo bem, e se você tem como mensurar uma métrica de sucesso, porque, lembrando que você foi contratada, né, para fazer esse algoritmo e vai, é, vender esse produto, então, é, até para mostrar que está tendo algum retorno legal, e está, é, com alguma acurácia interessante, não sei, aí você me diz o que você acha que era legal trazer, e se você considera ter feedback para as iterações, né, e ir ajustando o algoritmo de acordo com algum feedback humano ou não humano, depende do...

Ana: Testes, ai, eu não sei, os testes que eu tenho experiência seria fazer algum, tipo, trabalhar em cima dos requisitos e fazer um teste manual, assim, né, tipo, ai, então, a gente tem que procurar quais termos, ver como se comporta, é, ver se ele não está, tipo, enviesado, então, assim, se tiver escrito desenvolvedora, se ele não vai deixar de pegar, né, qual que é a importância que ele vai dar para isso, é, ter esses cuidados, e acho que o teste seria isso. Métrica de sucesso, não sei, não sei exatamente, teria que ver exatamente, tipo, ah, beleza, isso vai ser um NLP, e aí ele vai fazer aquelas buscas, e aí, se a gente tiver, tipo, um fuzzy match, eu vou pegar lá, é, sei lá, acima de 80

Entrevistadora: Sim, é legal. E aí, você considera feedbacks, para humanos, assim, para ter algum ajuste?

Ana: Ah, possivelmente, sim, né, se, então, é, sei lá, é, libera, por exemplo, um beta, deixa testando, então, o pessoal vai fuçando, coloca as coisas mais mirabolantes lá, vê o que retorna, coloca o simples, vê o que retorna, e aí, de acordo com o feedback, é que a gente vai corrigindo, né, o que tiver que ser corrigido.

Entrevistadora: Legal, e só uma questão, você falou sobre essa questão dos vieses, é, você acha que o algoritmo pode produzir algum tipo de viese, você falou do gênero, tem alguma coisa que você mensura, assim, logo de cara, que ele pode enviesar?

Ana: É, eu tenho essa preocupação, assim, de enviesar por conta da palavra no gênero feminino, né, então, assim, porque, sei lá, se tiver, vai procurar, e tem um monte de desenvolvedor e desenvolvedora, tem bem pouquinho lá, então, assim, qual que vai ser o retorno de desenvolvedora em relação ao desenvolvedor? Então, tem que ter um cuidado, um tratamento ali, para deixar tudo igual, pegar os desenvolvedoras também, e juntar tudo no desenvolvedor, e, sei lá, botar tudo em inglês, não sei, para poder, é, garantir ali que ele pegue, se for homem, se for mulher, né, tanto faz.

Entrevistadora: Tá legal. E aí, em relação, aí já é uma coisa mais abrangente, assim, os resultados e impactos, né, pensando na efetividade, você consegue pensar como avaliar, poderia avaliar essa efetividade, né, se as pessoas que estão sendo selecionadas para receber esse algoritmo, esse anúncio, desculpa, estão dando fit, né, com aquele anúncio de fato, e se você considera que o algoritmo em si, ele pode, da maneira que ele for formatado, ele pode gerar um impacto na contratação, né, assim, de pessoas de fato que vão ser contratadas, né, diante daquela configuração que você setou ali para o algoritmo.

Ana: A efetividade, acho que dá para ver em relação a, tipo, quantas pessoas que o algoritmo detectou que são um fit na vaga, de fato, se aplicaram à vaga, né, não sei se seria uma abordagem assim, tipo, aparece um anúncio, a pessoa clica, né, a gente seja isso, e aí a pessoa vai lá e se cadastra, então, talvez, no caso, comparar, tipo, quantas pessoas foram impactadas versus quantas pessoas clicaram, né, então, tipo, pegar essa métrica, né, de para quantas pessoas chegou o anúncio e elas só olharam e passaram, e quantas de fato clicaram e quantas de fato se cadastraram na vaga.

Na contratação, eu não sei, não sei, não, o que que poderia impactar no processo de contratação. Talvez facilitar, assim, né, tipo, uma pessoa que está procurando emprego, mas é aquela pessoa que não está ativamente procurando, a pessoa está insatisfeita com o trabalho, vai lá, então, pode ser que facilite, tipo, assim, ó, chegou aqui um negócio que é a sua cara, você está aí, infeliz no seu trabalho e tal, então, aqui, esse daqui é um fit interessante para você, talvez isso.

Entrevistadora: Tá legal. Foi isso, Ana, muito obrigada, foi bem legal, assim, você sabe bastante e você conseguiu trazer uma visão técnica muito legal. É, e acho que vai ajudar bastante aqui com todas as entrevistas que eu estou fazendo. Vou te colocar aqui o link desse formulário que eu falei no começo, né, são só alguns dados para você me passar, né, que acho que vai facilitar na hora de eu fazer a avaliação dos resultados, né, e muito obrigada, tá.

Eu queria te pedir mais uma coisa, se for possível, se você tiver outras mulheres, colegas, que têm um perfil que desenvolvem algoritmo, não precisa ser algum algoritmo específico, né, como você viu, assim, é só um exercício de pensamento mesmo, se você tiver contato e puder me passar, eu agradeço, tá.

Ana: A única que eu conheço, eu já falei com ela, ela falou que você já conversou com ela e marcou já para falar com ela, mas se eu pensar em mais alguma eu falo, sim.

Entrevistadora: Tá legal, obrigada.

Ana: Obrigada, tchau.

B.13

Denise

Entrevistadora: proposta aqui. É uma entrevista semi-estruturada, então tem alguns norteadores aqui pra gente ir discutindo. Vou até compartilhar a tela aqui, pra facilitar, que aí acho que você vai olhando, fica mais fácil de assimilar.

E a gente usa também, eu tô usando uma técnica que é do design, que a gente chama de Directed Storytelling, que é justamente, eu vou pedir pra você ir me dizendo e me contando qual o seu mapa mental, assim, né? Quais são as suas decisões pra uma ação específica que eu vou te passar, né? Vou contextualizar, e aí vou passar um cenário e pedir pra você ir me falando, assim, o que você faria. Mas a gente vai por etapas aqui acompanhando esses norteadores, tá bom?

Denise: Tá bom.

Entrevistadora: Primeiro eu queria saber, pedir pra você dar uma contextualizada em relação à sua experiência com algoritmos, sua experiência na parte de dados também.

Não precisa falar de projetos específicos, né? Mas assim, mais pra gente entender exatamente como você se insere.

Denise: Tá. Então, eu trabalho na área de reservas e reservatórios desde que eu ingressei na Petrobras, e eu acho que foi em 2017 ou 2018.

Bom, eu trabalhei um tempo com uma parte mais de engenharia de reservatórios, estudar projetos específicos, né? Estudar comportamento, previsão de produção de óleo, principalmente de projetos específicos. E aí, em 2017, 2018, comecei a trabalhar num programa mais estratégico, assim, que deveria ter uma visão mais macro do conjunto de previsão de produção de óleo da empresa toda, assim. E aí foi aí que eu comecei a trabalhar com dados.

Nessa época, sem muita experiência com algoritmo, mas com análise de dados, assim. E aí, bom, e aí, a partir disso, isso me forçou, de certa forma, porque o volume de dados era muito grande, me forçou, de certa forma, a começar a trabalhar com técnicas, assim, das mais simples às mais complexas de análise de dados. Então, esse algoritmo de regressão, essas coisas mais simples, assim, para identificar padrão, identificar outlier, enfim.

E aí, depois, por causa de uma mestrada, isso foi uma virada, assim. Nessa época, dentro da mesma equipe, tinha o pessoal trabalhando com uma temática que não tem exatamente a ver com a anterior, mas trabalhando com semântica, que é... Enfim, a gente estava trabalhando em tudo o que era habilitador. Eles

estavam trabalhando, no caso, com tudo que era habilitador para soluções conservativas em termos de óleo e gas, assim, em língua portuguesa.

Então, esse que era um negócio que tinha começado muito antes, começaram levantando o levantando o conjunto de dados que a gente teria de texto para óleo e gas em língua portuguesa, depois a gente começou a trabalhar algoritmos, tal, e eu entrei numa fase, agora, de modelos de processamento em linguagem natural diversos, assim, modelo de reconhecimento de entidade nomeada, modelo de ATC link, modelos de processamento de linguagem natural. Essa é a minha experiência com algoritmos, mais recentemente eu tô menos em algoritmo, menos treinando, montando algoritmos, foi uma coisa que eu nunca era muito boa, tem pessoal da informática que é melhor nisso, mas mais inteligência do negócio para ajudar quem está montando os algoritmos, assim. Então, avaliar, validar, levantar métrica, fazer a análise crítica, porque que ficou bom, porque que não ficou bom.

Essa é a minha experiência.

Entrevistadora: Ótimo, muito bom, Denise. Eu vou passar, então, para você o contexto aqui desse cenário fictício, e aí vou pedir, assim, claro, quando tiver uma dúvida pode perguntar aqui, mas eu vou pedir para você ir dizendo, assim, o que você faria.

Você pode falar, aprofundar em técnicas, se você preferir, mas a minha ideia mesmo é entender, assim, quais são os caminhos para sua decisão, suas tomadas de decisão. Então, imagina que uma empresa de tecnologia de porte médio contrata uma empresa de divulgação de vagas, para anunciar uma oportunidade. Então, você trabalha nessa empresa de divulgação de vagas e você vai ter que fazer um algoritmo para selecionar as pessoas que vão receber a oportunidade.

Então, imagina, assim, esses anúncios que aparecem em Facebook, Instagram, que já tem um certo recorte, né, assim, já foi lido o seu perfil ali, então, por isso, esses anúncios estão sendo direcionados. Então, a ideia é exatamente essa, você desenvolver um algoritmo para alguns requisitos, né, importantes para selecionar pessoas que podem dar fit com essa oportunidade. Então, a empresa, ela passa esses requisitos, assim, mais abrangentes.

É uma vaga de tecnologia para um profissional pleno, com skills de desenvolvedora ou desenvolvedora full stack, com formação em engenharia de software, tecnologia da informação, sistemas e afins, né, mais de áreas de ciências e exatas. Em geral, experiência de cinco anos e vai ser um regime remoto, tá? Então, eu vou para a primeira parte, depois eu volto para esse slide para a gente lembrar, deixar claro aqui os requisitos. Na parte de planejamento e concepção dessa solução, em relação à definição dos requisitos, como você

refina a definição desses requisitos, como você define, né, para o algoritmo realmente, especificamente, quais seriam esses requisitos e quais dados você utilizaria, né? Você não precisa ficar preocupada, por exemplo, pegaria no LinkedIn, mas não sei se eu teria acesso.

Vamos fingir que está tudo aberto, que está tudo fácil de pegar, né? Um mundo fantasioso aí. E seriam, assim, quais os dados seriam importantes você selecionar, utilizar para fazer o modelo, né, que a gente falou de direcionar para outras pessoas esse anúncio específico.

Denise: Tá bom.

Bom, primeira coisa para definir os requisitos, eu definiria o que eu quero de saída desse algoritmo, que eu imagino que seja um conjunto de pessoas, né, não é a pessoa, enfim, um conjunto de pessoas de interesse e acompanhado desse conjunto de pessoas, talvez uma escala do meu interesse, assim, que aí, enfim, não importa como que eu vou calcular muito essa escala, mas talvez adequação à vaga, né, que eu tô publicando, e eu colocaria também os requisitos da vaga que eu tô definindo. Então, procurar por dados, procurar por desenvolvedores Full-stack, vou supor aqui que eu tenho um banco de currículos, ou que eu tenho que eu tenho um banco de alguma empresa de RH, com pessoas pré-cadastradas. Então, por exemplo, não selecionaria as pessoas que estão se candidatando a serem enfermeiras, né? Eu colocaria esses critérios como os critérios para calcular esse índice, assim, então, qual é a titularidade da vaga que aquela pessoa se candidatou, ou se ela já trabalha como profissional nesta vaga, e vagas afins que estejam, e aí, isso porque eu trabalho com linguagem, talvez, que estejam no, sei lá, num espaço vetorial similar, assim, né, que sejam temas afins com o desenvolvedor Full-stack, porque vai ter muita coisa que é análogo.

Então, eu também colocaria como demanda desse algoritmo ter um match na formação desses temas que me interessem, e eu colocaria como critério, eu tô entendendo que experiência não é um critério flexível também, né? Eu colocaria como, talvez, como corte, assim, se tiver menos de cinco anos de experiência, é um sim ou não, assim, uma árvore de decisão, né? Não selecionar. Ou, se isso for mais flexível, isso compor o meu índice de adequação à vaga, né? Se tiver menos de cinco anos de experiência, a gente penaliza um pouco esse, esse, essa pessoa. Aí, eu tô entendendo que regime remoto, a questão da localidade não vai ser um critério pra mim, mas uma coisa que eu não sei se faria parte nessa etapa é avaliar se existem outras, é, bom, as características listadas são essas, né, mas se existem outras características que a empresa valoriza, enfim, que podem ou não, é, posicionar melhor esse profissional que eu tô, que eu tô targeteando, né?

Entrevistadora: Tá legal.

Denise: É isso.

Entrevistadora: E aí, é, em relação à escolha do método, acho que você já falou um pouco, mas, assim, se você consegue, consegue dizer se seria uma técnica de machine learning, alguma coisa, né? Você falou de NLP, mas, é, como, se essa escolha do método também tá relacionada a algum desafio técnico que você pode prever, pode mensurar, é, em relação a essa integração, né, desse algoritmo?

Denise: Essa é a parte mais, que eu, engraçado, essa é a parte que eu consideraria mais difícil, assim, porque, vamos lá, a gente pode fazer, porque eu faria isso, eu tentaria uma técnica que atribui pontos, assim, que atribui, que coloque essas pessoas numa escala de pontos em função, é, em função da, do atendimento a determinados requisitos, é, mas isso é pouco machine learning, assim, é quase o método, é, é quase o método analítico, assim, né, de ranqueamento, é, então, acho que é isso, se eu partiria até de um método simples, se esse método simples é ruim, não me atende, talvez, é, avaliar com base no, nos dados dos colaboradores que eu tenho dentro da minha empresa, aqueles, aqueles que seriam adequados para essa vaga, e considerar isso como um conjunto de dados anotados, né, então, assim, o fulaninho ocuparia o primeiro lugar, o ciclâninho, e aí treinar um, treinar uma, uma técnica qualquer, ou, talvez, uma regressão, ou uma técnica mais simples, assim, que, é, que usa esses dados anotados dos meus próprios colaboradores, apesar disso poder ser pouco dado. Eu acho que isso, desafio técnico é, desafio técnico é, assim, é, validar essas coisas, porque, bom, sendo uma, um algoritmo desenvolvido por uma empresa de RH, você sempre precisa de gente do negócio para validar se o algoritmo está desempenhando bem, né, de gente, né, o difícil, pelo menos esses são os desafios técnicos que eu enfrento sempre, assim, é gente, é, é, e, principalmente no cenário que você está considerando que dados disponíveis não são um problema, né, se os dados disponíveis não são um problema, a dificuldade é gente, assim, para fazer uma apreciação do, do resultado do seu algoritmo.

E como o algoritmo seria integrado ao sistema de divulgação de vagas da empresa, ah, isso é muito interessante, eu acho que o ideal seria, seria LinkedIn, né, seria um algoritmo que, é, pensando em divulgação de vagas, assim, seria um algoritmo que captura informações de LinkedIn, e, enfim, talvez dispare, dispare uma mensagem ou dispare a oportunidade da sua vaga, dê o anúncio da vaga para esses usuários específicos, né. É, não sei se é empresa que tem algum outro tipo de divulgação de vagas, mas, enfim, essa parte de vagas não é muito a minha especialidade, porque eu trabalhei no mesmo lugar há 12 anos.

Entrevistadora: Tá legal, mas respondeu bem aqui.

E aí a gente, é, em relação aos testes e validação, né, pensando se você tem alguma ideia de teste, de método de teste para realizar, para ver, né, balizar ali o resultado que está saindo do modelo, é, alguma métrica de sucesso, eu acho que você já até trouxe alguma coisa também, mas pensando no fato de que você está desenvolvendo um produto que vai ter um resultado, né, e aí você está vendendo esse produto, né, e feedbacks, como você falou, né, e para considerar iterações nesse algoritmo, com, sendo interno ou externo, humano ou não, para ir melhorando nesse algoritmo.

Denise: Beleza, então, qual tipo de teste para realizar, é o que eu comentei, eu acho que eu faria testes, é, ou criando perfis do usuário ideal, ou coletando, coletando, usuário não, do selecionado da vaga ideal, ou a partir de pessoas reais, que eu considero que seriam ideais para vagas e que eu conheça, para ver se o algoritmo tem o mesmo entendimento que eu tenho, né, que aquela pessoa é ideal. É, mas, isso teste, mas métrica de sucesso, é, eu imagino que a métrica de sucesso seria, é, o algoritmo ser capaz de trazer um conjunto de pessoas adequado, assim, uma quantidade de pessoas adequada para esse meu processo, é, com, aí é difícil, porque você tem que considerar que essas pessoas estão disponíveis no mercado, mas considerando que elas estão disponíveis, então, o algoritmo trazer um conjunto de pessoas adequado, com uma métrica que as diferencie de maneira interessante, para eu poder conduzir as partes não tão automatizadas do processo, né, então, por exemplo, se o algoritmo me trouxer dois mil candidatos para uma vaga, eu não vou considerar ele adequado, não é uma métrica de sucesso, vai ser muito difícil, é, vai ser muito difícil eu conduzir entrevista, enfim, com tanta pessoa, mas se for, é, se trouxer dois fulanos, aí, também não é muito bom, né, acho que essa seria uma métrica de sucesso, o conjunto de pessoas, assim.

E avaliar feedback, é, eu tenho pouquíssima experiência, assim, com feedback, mas o que a gente, experiência pequena que eu tenho e que funciona, assim, é um feedback do tipo, é, as cinco estrelas, assim, como você avalia, então, é, consultar a empresa de RH que está conduzindo o processo, é, para cada, para cada pessoa, talvez, para cada perfil, para cada pessoa perfilada, assim, se está bom ou não, né, para poder retroalimentar o algoritmo, eu faria feedback simples, não muito, não muito complicado.

Entrevistadora: Tá legal. E aí, por fim, Denise, em relação à efetividade e impacto, né, assim, aí é uma opinião sua, né, se você considera que é efetivo usar esse tipo de algoritmo para essa ação específica, né, fazer essa, esse filtro inicial, e se você considera que isso gera um impacto na contratação, né, você considerando que são pessoas e que você vai já fazer um nicho ali, né, fazer um

filtro de pessoas que vão receber essa oportunidade.

Denise: Essa é uma, bom, a efetividade, eu acho que a efetividade pode ser muito alta, é, no sentido de diminuir o tempo buscando manualmente dentro desses dados, então, considerar que eu treinei um modelo com dados e que automaticamente esse modelo me diz se aquele candidato é bom ou não, muito mais simples do que ficar fazendo uma avaliação humana. Por outro lado, o impacto, é, pode, eu acho que gera um impacto significativo, e aí falando, não do tamanho do impacto, mas qualificando esse impacto, Lari, é uma coisa que pode ser, é, pode ter impacto tanto positivo quanto não tão positivo, assim, do ponto de vista de fator humano da empresa, assim. Uma coisa que eu acho muito positiva, é, de algoritmo, é a questão da anonimização, dele, é, a gente tem o viés de decisão quando você olha para pessoas, né, isso eu vejo muito na minha empresa para processo seletivo interno, por exemplo.

Se é uma ou outra pessoa que está candidato e que tem um perfil parecido, ele é sujeito ao viés do contratador em relação aos fatores de idade, gênero, enfim, não necessariamente a capacitação técnica, então, se o algoritmo tiver, se atendo à capacitação daquela pessoa e as características que independem, né, que independem desses fatores, é um impacto positivo, você pode trazer, é, você pode trazer uma diversidade maior, você pode acabar se deparando com candidatos que você não imaginaria, que você não selecionaria a priori.

Por outro lado, algoritmos, algoritmos não são isentos de vieses também, porque quem define o que é uma métrica de sucesso, quem define o que vai ser avaliado, são os humanos, então, se você define a priori, é, características que vão favorecer o outro grupo, você pode usar esse algoritmo no sentido de, do contrário, de diminuir a diversidade da sua empresa, outra vez, de trazer perfis, é, que não correspondem exatamente à cultura da empresa, ou que não são adequados, enfim, se você for muito para o lado técnico, pode, você pode perder, de repente, as habilidades, outras habilidades, né, que você gostaria, talvez até em detrimento um pouco dos cinco anos de experiência, por exemplo, você acha que tem impacto significativo, ele pode ser usado, é, ele pode ter um efeito qualificando, né, o que eu acho interessante, no sentido de aumentar a diversidade, quanto de reforçar, de reforçar vieses, a depender do que, depender dos requisitos, então.

Entrevistadora: E você considera, Denise, que seria interessante ter, tipo, um checklist, sabe, para alguma, para essas soluções que, que impactam, né, socialmente, que, é difícil dizer quando um algoritmo não impacta, né, a vida das pessoas e as coisas no dia a dia, enfim, mas, é, ou que tem alguma tangibilidade com essas questões sociais, você acredita que se tivesse um checklist mais abrangente para esses algoritmos específicos, né, que vão ter

essa ligação direta ali, é, com a população, né, com o nicho de pessoas, se tivesse um checklist para garantir um pensamento mais inclusivo e diverso, seria interessante, no ponto de vista de sistematizar esse processo?

Denise: É, cara, super complexo, porque a gente tá falando de inteligência de negócio de uma empresa específica, assim, o ideal, nesses casos, é, não necessariamente o checklist, mas um, ou um software aberto, ou um software, é, revisado por, tem empresas que trabalham com, é, como eu posso dizer isso, nomear o que elas fazem, assim, com, é, uma justiça, é, no ambiente dos algoritmos, então, você revisar isso por uma, por uma empresa especializada, por exemplo, tem muita gente que discute, não é um processo simples, eu acho, né, não sei, assim, e também, algoritmo aberto, você tá entregando, não é uma coisa muito popular, assim, entre as empresas, que você entrega a inteligência do negócio, né? É, mas acho que, talvez, não um checklist, mas um processo de verificação por um terceiro, assim, uma ONG, uma universidade terceira, com grupo de pesquisa, seria, é um processo legal.

Entrevistadora: Legal, Denise, muito obrigada, é, foi ótimo, assim, você tem uma visão muito boa de, né, que é interessante, agora, já falando offline, aqui, da entrevista, né?

Denise: Aham.

Entrevistadora: Mas, você tem uma visão estratégica muito maneira, que é diferente, né, da, às vezes, do cara que tá muito no dia a dia do algoritmo, né, e você já pensa lá, já começou a pensar na finalidade e tal, falei, gente, mas muito bom.

Denise: Ah, que legal, tem um, eu ia até recomendar um livro, que eu mesma não li, e não tá aqui na minha cabeceira, agora, que é uma pesquisadora, na verdade, agora, eu não sei se eu, eu não sei qual gênero, que essa pessoa tá se identificando, que ela tá em transição de gênero, é, mas ela não publicou nada a respeito, então, vou falar, que é uma pesquisadora da PUC, que agora tá na Unicamp, trabalhou em GT de transição, de ser esse tecnologia do governo, chama Nina Daora, e ela escreveu um livro que chama Algoritmos, inclusive, eu comprei esse livrinho, e ela fala muito, assim, de, ela se denomina como hacker antirracista, e ela e umas gringas, aí, que são ótimas, assim, fazem esse estudo, assim, sobre essa coisa do algoritmo não ser uma, do algoritmo não ser isento de viés, assim, que, pra mim, foi uma coisa meio reveladora, assim, que eu nunca tinha pensado.

Entrevistadora: É, eu sigo ela no Instagram, mas nunca li o livro dela mesmo, e vou, vou dar uma aprofundada, bem interessante, ela traz, realmente, pautas muito boas.

Denise: Eu comprei o livrinho dela, mas tá aí, embalado, junto com 200

outros livros.

Entrevistadora: Tem bem como é que é, tem muito como que é. Fica bom na prateleira, né?

Denise: Ai, tá bom.

Entrevistadora: Denise, muito obrigada, quando você puder, eu mandei aqui no chat o nosso link, tá? O link do formulário, mas fica à vontade, não sei se é agora, e tô te devendo uma cerveja.

Denise: Sim, vou querer, vou estar querendo. Vamos.

Entrevistadora: Beijo, Denise, bom feriado pra você.

Denise: Beijo, bom feriado. Beijão, tchau.

B.14

Joana

Entrevistadora: Porque isso, às vezes, modifica, tá certo. Tá bom. Eu vou compartilhar a tela com você, mas como eu falei assim, a ideia é só para ter alguma referência visual ali para você ir se guiando.

Esse método que a gente usa em cima da entrevista semi-estruturada chama Directed Storytelling, que é você realmente ir falando o que você está pensando para realizar alguma atividade, tá bom? Se você precisar pegar a tela para mostrar alguma coisa, alguma ferramenta, plataforma que você acha que vale a pena, fica à vontade, tá bom? Tá vendo a minha tela? Eu vou projetar, vê se foi projetando também. Então, vamos começar. Primeiro, eu queria saber um pouco se você pode contextualizar sua experiência com algoritmos e um pouco da sua experiência profissional também.

De repente, até se você quiser falar da sua experiência na empresa hoje.

Joana: Hoje eu atuo como cientista de dados sênior numa consultoria, então a gente presta serviços de analíticas para outras empresas, principalmente do ramo de mineração. Eu estou dentro da área de preditiva, que a gente tem também outra área de prescritiva, otimização e também TI, mas eu estou nessa parte de preditiva, onde a gente desenvolve alguns modelos ou análises, dashboards, coisas desse tipo.

Minha experiência anterior foi numa empresa de serviços financeiros, onde eu também atuava como cientista de dados. Eu era de uma área central, uma área de dados, e atendia outras áreas da empresa com diversos produtos. Antes disso, eu trabalhei numa seguradora, que era algo mais voltado para análise de dados, mais validação, trabalhava pouco com modelagem.

Antes disso, fui numa empresa de telecomunicações, onde eu construía algumas químias para extrair relatórios, coisas desse tipo. E aí, minha expe-

riência com algoritmos. Bom, já desenvolvi modelos supervisionados e não supervisionados.

A metodologia depende muito do projeto, mas já desenvolvi modelos de regressão, de classificação, que eu desenvolvi bastante esse tipo, e também alguns de segmentação. Para desenvolver os modelos, quer chegar lá?

Entrevistadora: Não, não sei se você quer trazer uma questão específica de alguma experiência que você teve, mas aí para desenvolver a gente já passa para a próxima parte, para você ir descrevendo. Então, eu vou te falar de um cenário fictício, tá? E aí a gente vai destrinchando isso por fases.

Imagina que uma empresa de tecnologia, de porte médio, te contrata, contrata uma empresa de divulgação de vagas, vamos dizer que você trabalha nessa empresa de divulgação de vagas, para anunciar a seguinte oportunidade. É uma vaga de tecnologia para profissional pleno. As skills são correspondentes a um desenvolvedor ou desenvolvedora full stack, com formação em engenharia de software, tecnologia da informação, sistemas, estatística e ciências em geral.

Uma experiência de aproximadamente cinco anos para trabalhar em regime remoto, tá? Então, essa empresa traz apenas esses requisitos, entrega, né, para você, vamos dizer assim, e aí você precisa desenvolver um anúncio. E como seria esse anúncio? Você sabe quando você tá no Facebook, né, a gente, antigo Facebook, apareciam aquelas propagandas, né, algumas coisas assim que te puxam, né, que são propagandas mesmo, que tem a ver com o seu perfil, né?

Então, aquilo ali tá aparecendo, aquele conteúdo, a propaganda tá aparecendo justamente porque algum algoritmo, né, foi feito ali, leu alguma coisa dos seus dados, e aí eles consideram que aquilo tem a ver com o seu perfil. Então, a ideia é você pensar nisso, assim, não precisa se ater especificamente ao tipo de plataforma, qual plataforma, se você achar que funciona, por exemplo, pensar no LinkedIn, alguma coisa assim, mas já consolidada, pode pensar, mas a ideia mesmo para você, porque a gente quer entender, para você dizer, né, nos dizer como que você vai estruturar esse desafio, né?

Joana: Então... E aí você quer só o headline, assim, a chamada ou todo o anúncio?

Entrevistadora: É, o que eu vou pedir para você é ir pensando, assim, em etapas, por exemplo, no começo, né, como você faria esse planejamento, concepção, então você... Como seria esse refinamento de definição de requisitos, né, que tem aqueles requisitos ali bem simples e bem diretos, e como você faz a extração desses dados, quais dados seriam utilizados? A princípio isso, aí depois a gente vai falando por partes também.

Joana: Tá, mas aqui a gente está selecionando currículos, é isso?

Entrevistadora: Isso, a gente está, na verdade, selecionando pessoas que

tenham o perfil técnico que podem preencher a vaga que você vai anunciar, tá? Então, você tem uma vaga, vai desenvolver, né, depois um anúncio, e aquele anúncio ele vai aparecer para quem você acha, você considerou, que tem o perfil daquela vaga, tá? É tipo a gente está navegando no Instagram, e aí começam a aparecer aquelas propagandas no meio ali, que tem a ver com seu perfil, porque você falou alguma coisa e tal. É um pouco isso, mas, assim, você vai primeiro rastrear os dados dessas pessoas, né, os dados profissionais, aí de repente até realmente usar um LinkedIn pode ser importante, né, assim, pode ser que case. E aí, assim, que dados que você pegaria dessa pessoa para você conseguir cumprir esses requisitos aqui, iniciais, né?

Joana: Tá, bom, primeiro eu focaria um pouco nas hard skills, que eu acho que ajudaria a filtrar ali os candidatos.

Então, todas as linguagens que forem importantes para essa vaga, que aí eu não sei especificamente com o que esse desenvolvedor full stack trabalha, mas todas essas linguagens, e também a formação nesse tipo de área, né? Então, tentaria buscar por essas palavras. Ano passado eu fiz um curso de redes neurais, de, na verdade, aprendizado, aprendizagem profunda, e aí focava bastante redes neurais. Então, talvez eu tentaria fazer um NLP para ver palavras que se aproximassem também, né? Porque às vezes só tecnologia da informação, só fazer um filtro, talvez eu não encontre essa, especificamente essas palavras, né? Então, tentaria ver palavras próximas.

Talvez fazer algum filtro pela idade, para buscar mais ou menos um profissional que tenha cinco anos de experiência, ou pegar a partir da data de formação, né, ou dos seus primeiros desses primeiras experiências profissionais, para ver alguém que tenha pelo menos cinco anos de experiência.

Entrevistadora: Tá legal. Então, assim, teria que fazer uma definição melhor dos requisitos, de repente, com a empresa, e para, até como você falou, né? Assim, não sei exatamente quais são os hard skills que tem a ver, né? Qual é esse perfil, e você vai buscando ali em relação à formação, experiência, né? Fazendo esse filtro de dados ali.

Informação. Tá bom. Aí, em relação ao desenvolvimento do algoritmo, você já até falou alguma coisa de método, mas você teria como explicar um pouco mais, ou se você usaria métodos combinados, não sei, e se tem alguns desafios técnicos que você prevê que podem ser endereçados com a escolha do método, né? Assim, para você já pensar, assim, de acordo com esse desafio, vou escolher tal método, e como você faria essa integração e implementação, de fato, do algoritmo?

Joana: Eu acho que teria algumas saídas, assim, talvez fazer uma segmentação para buscar perfis próximos, talvez usar alguém que já esteja na empresa,

que se aproxime do perfil, porque a gente não tem uma base de dados, assim, com label, né? De quem é um potencial desenvolvedor ou não.

Então, acho que talvez eu partiria para essa parte da segmentação, e aí, para tratar esses dados, né? Como a gente está falando de dados não estruturados, aí talvez para isso utilizaria alguma técnica de encoding para buscar ali o valor semântico dessas palavras próximas do que a gente está procurando.

Entrevistadora: Tá legal. E aí para teste e validação? Qual tipo de teste você realizaria? E você consegue definir também alguma métrica de sucesso, né? Considerando que você foi contratada para fazer esse algoritmo, né? Como é que você também relata para a empresa que te contratou, que está funcionando, que está sendo efetiva, né? Essa definição mesmo das métricas de sucesso, e se você considera ter algum feedback para as iterações, né? Feedback humano, não humano, quais seriam?

Joana: Como a gente está, como eu decidi usar um algoritmo de segmentação, né? Acho que seria um pouco difícil avaliar esse resultado sem utilizar esse feedback.

Então, acho que seria importante ter esse feedback dos entrevistadores para verem se os perfis estão correspondendo aos pré-requisitos da vaga, ou se teve algum gap que eles identificaram que a gente pudesse melhorar e talvez aperfeiçoar esse algoritmo, dando, talvez, usando outras variáveis ou pegando outras informações, talvez tirando alguma coisa desse tipo.

Entrevistadora: Tá legal. E aí, assim, alguma métrica você acha que consegue definir?

Joana: Assim, acho que a gente consegue avaliar para uma curar-se a partir desse, a partir do feedback dos entrevistadores, eu acho.

Entrevistadora: Tá legal. E em relação aos resultados e impactos, né? Assim, a efetividade mesmo, né? Do algoritmo, se for algo do algoritmo na recomendação dos candidatos, aí a gente já falou um pouco antes, né? E o impacto na contratação de uma maneira geral, assim. Como a gente está falando de pessoas, né, que vão ser selecionadas a princípio pelo algoritmo para que a presente, que seja apresentado esse anúncio, né? E aí, por ventura, algumas dessas pessoas podem vir a ser contratadas.

Isso é um impacto social, vamos dizer assim, né? Você acha que seu algoritmo, né? O algoritmo que você vai produzir, tem essa relação direta, né? No impacto em relação à contratação, ao desenho dessas pessoas que vão contratar, ao perfil, enfim.

Joana: É, acho que impacta bastante, né? Porque talvez pessoas com experiências diferentes não vão entrar nessa filtragem. Então, é algo que tem

que tomar um certo cuidado para a gente não ficar muito específico num grupo de pessoas, né? Mas eu vejo que a efetividade está em facilitar a busca por esses candidatos, né? E ser mais assertivo no público que a gente está buscando.

Então, ao invés de colocar esse anúncio para todo o LinkedIn, onde vai ter pessoas que trabalham com RH e marketing, outras áreas, a gente vai focar em pessoas que trabalham com desenvolvimento. A gente vai reduzir custo, vai buscar pessoas que realmente tenham esse perfil.

Entrevistadora: Quando você falou que tem que ter um cuidado, né? Para fazer esse recorte, que cuidado você acha que criaria um alerta, assim, para você?

Joana: É, sim. Eu vejo que algumas ferramentas de recrutamento que existem no mercado, elas filtram muito por algumas skills específicas, assim, eu acho que alguns candidatos acabam ficando de fora. Então, eu acho que é importante ter um cuidado para não só filtrar, mas talvez relacionar, assim, com as principais, os principais requisitos ali, mas não necessariamente alguém que tivesse todos aqueles hard skills, porque, bom, a pessoa pode não ter colocado ali, ou ter que verificar a relevância, também, de determinado hard skill para a vaga. Então, eu vejo essa preocupação, assim.

Entrevistadora: Tá legal. Foi isso mesmo, Joana. Foi ótimo, obrigada. Deu para ter uma ideia, mais ou menos, do seu mapa mental, né? Assim, como você iria destrinchando. Vou enviar para você, já vou enviar por aqui, que aí fica mais fácil, o formulário, tá? Quando você puder, sem pressa, não precisa ser agora, tá? Quando você puder responder, deixa eu enviar aqui para você, mas já vai ser ótimo. E te pedir, não sei, eu sei que é chato, mas se você tiver, principalmente mulheres, para indicar, tá? Porque eu estou entrevistando a mesma quantidade de homens e de mulheres, e a gente sabe que é mais difícil, né, conseguir mulheres com esse perfil, né, técnico.

Se você tiver para indicar, eu agradeço, tá?

Joana: Tá. Vou falar com algumas pessoas e eu te informo. Entrevistadora:

Tá ótimo, muito obrigada. Boa semana para você. Obrigada. Tchau, tchau.

Joana: E você também, tchau, tchau.

B.15

Leandra

Entrevistadora: A transcrição tá em português? Acho que tá, tá certo. Eu vou compartilhar a tela aqui para a gente ir guiando uma certa pauta, assim, dessa entrevista, mas é uma entrevista semi-estruturada com o método que a gente usa no design, que a gente chama de Directed Storytelling, que

consiste basicamente, assim, eu peço para você fazer alguma ação, né, e você vai me narrando, assim, o que você pensa, como você faria, o que você, sei lá, uma ferramenta, algum método, questões, assim, que vierem à cabeça. Então, não tem problema se depois você quiser mudar de opinião ou não tem certeza, vai ter que pesquisar, enfim, fica tranquilo, tá?

Leandra: Tudo bem, beleza.

Entrevistadora: Se você precisar compartilhar sua tela, fica à vontade, tá? Assim, não sei se você vai querer, se você quiser mostrar alguma coisa, né, algum... deixa eu voltar para longe. Alguma ferramenta, tá bom?

Leandra: Tá, tranquilo.

Entrevistadora: Eu vou começar pedindo para você contextualizar um pouco melhor a sua experiência profissional, acadêmica, que você tiver, com algoritmos, né, e um pouco do que você faz agora tá entrando, né, imagino que ainda esteja se ambientando, mas só para eu entender o contexto geral da sua experiência.

Leandra: Tá, claro. Então, eu tenho profissional em Telecomunicações. Então, eu tenho formação técnica de Telecom.

Depois entrei, né, parte profissional, eu trabalhei em empresa de Telecomunicações também, como técnica, mas eu também fiz uma transição para computação. Então, primeiro eu alterei, né, a minha graduação para Sistemas de Computação e depois troquei também de trabalho, né, fui para uma empresa de tecnologia, consultoria também, né, assim como a Deloitte, mas não era a Deloitte. E tanto, né, na graduação, né, de Sistemas de Computação, eu tive ali contato com quais matérias, né, de fundamentos de programação, Python, Java.

Então, tinha ali bastante matérias, né, envolvendo programar em si, enfim, você escrever, né, um algoritmo, fazendo os exercícios, né, que envolvem escrever e programas. E no trabalho eu entrei numa equipe que mexia com Big Data, fazia projetos, né, de Big Data e depois tive também experiência com BI. Então, nos projetos que eu passei, eu via muitos scripts, né, muitos programas para trabalhar ali com extração de dados, com transformação e armazenar, né, o básico ali de ser o ETL, né? Então, era na linguagem de Python, cheguei a ver alguma coisa em escala, mas com essa finalidade, né, para serem sistemas de informação.

Então, aí aqui, mais recentemente, né, então eu me formei na graduação em 2022 e aqui na Deloitte, né, entrei faz seis semanas e eu tô no, é, eu tô no, né, cargo me fala, né, né, cargo, de Engenheira de Dados, né, que foi o que eu estava me especializando, né, na outra empresa e aí aqui teve essa oportunidade, né, então eu tô dando continuidade. E aí a minha, a minha

experiência, né, com a questão de algoritmos é o mais recente, né, nos últimos quatro anos, tem sido voltados para essa área de Engenharia de Dados.

Entrevistadora: Legal.

Leandra: É isso.

Entrevistadora: Muito bom. Então, eu vou explicar um pouco da simulação, né, desse cenário fictício, para você me dizer o que que você, quais seriam as tomadas de decisão, né? Então, imagina que uma empresa de tecnologia de porte médio contrata uma empresa que você trabalha para fazer uma divulgação de vagas, de um anúncio de, de vaga, né, de vaga de emprego, né? Então, eles passam apenas esses requisitos aqui.

É uma vaga de tecnologia para um profissional pleno, com skills para um desenvolvedor ou uma desenvolvedora full-stack, com uma formação de engenharia de software, tecnologia da informação, sistemas, afins, uma experiência de cinco anos, aproximadamente, desculpa, para trabalhar no regime remoto. Então, só para explicar melhor, essa empresa de tecnologia de porte médio, ela tá com uma vaga aberta e só que ela quer anunciar essa vaga para alcançar as pessoas que tenham um perfil específico para ocupar essa vaga, né? Estão querendo acelerar ali a seleção desses candidatos, já fazendo um certo recorte no início para já virem pessoas, né, que têm de uma certa forma algum fit, né, skills e questões que vão dar um fit ali com esse anúncio, com esses requisitos específicos que a empresa está passando.

Leandra: Tá, no caso, só para confirmar o entendimento, eu já estaria contratada, né, nessa empresa de divulgação, eu já sou dessa empresa de divulgação, e no caso, a empresa de tecnologia que contratou a minha empresa vai ser para anunciar uma vaga, né, e com esses tópicos

Entrevistadora: Isso, a empresa que te contratou, né, contratou sua empresa, ela quer que vocês produzam um anúncio, mas que tenha um algoritmo ali por trás que vai já selecionar as pessoas que eles, que vocês entendem que cabem para essa vaga, né? Então, assim, já de repente tem que direcionar e ler um pouco o perfil de quais pessoas que vão ser selecionadas para ver, né, para aparecer esse anúncio.

Tá? Então, o primeiro passo seria pensar, assim, no planejamento e na concepção, né? Primeiro na definição dos requisitos, como você define melhor, ou se você, em cima dos requisitos que foram passados, como que você pode fazer um refino desses requisitos, e quais dados você utilizaria, como você conseguiria. E essa parte técnica, vamos dizer assim, por exemplo, se você for extrair do LinkedIn, você não sabe se pode uma coisa ou outra, não precisa ser ateu, tá? Assim, vamos dizer que a gente está no mundo ideal, tudo aberto, né? Assim, sem nenhuma restrição.

Leandra: Certo, entendi.

Tá, pensar como você definiria os requisitos para o algoritmo, certo? Então, vai além de simplesmente divulgar, né? E até, pois é, usando o LinkedIn, vai além de criar uma publicação com aqueles tópicos, né? Seria realmente algo que captasse, né? E já selecionasse também? Seria... Ou, de repente, isso já é uma parte do requisito, né? Não sei.

Entrevistadora: É, não, por exemplo, a gente está no Instagram, aí aparece, sei lá, você falou que você quer comprar um sapato, aí começa a aparecer publicação de sapato de alguma forma, você entrou ali na bolha daquelas pessoas que podem ser que sejam interessadas no sapato, né? Aí a ideia é, assim, o algoritmo que vocês vão criar, né, da sua empresa, vai capturar pessoas que podem ter interesse nessa vaga específica, que tenha competência, pelo menos, assim, aparentemente, né, para ocupar essa vaga.

Leandra: Certo, é porque... Então, já estava vendo as ideias na cabeça, no sentido, assim, de que, desse exemplo que você falou, eu vou, a gente tem essa linha, né, de divulgar a informação e, dependendo de quem está ali a acessar e, por exemplo, se inscreva aqui, né, dependendo, então, de quem se inscrever, seria um tipo de linha para dar e filtrar os candidatos e tal.

Mas também vem na minha cabeça aqueles sites que existem hoje em dia, né, de que você, por exemplo, tem o seu perfil lá e, aí, as vagas, né, que são cadastradas já podem fazer um filtro, né, no banco de dados que já está cadastrado dentro do site.

Entrevistadora: Isso, isso, assim, pode ser desse jeito também, assim, né, as vagas já vão ser direcionadas.

Leandra: Tá.

É, assim, eu acho que eu botaria isso como requisito, né, de ter essa definição de... Acho que também isso esbarra, né, sobre a questão dos dados utilizados, né, como que, como, é, como, como que seriam os dados de entrada, né, se vai ser, se é algo mais, depende do candidato, né, que veria ali o anúncio da vaga para se inscrever ou se, a partir de um banco de dados já cadastrado, você já faria os filtros, o algoritmo faria os filtros para selecionar. Mas acho que eu vou, então, partir desse segundo ponto, que aí já, na minha cabeça, acho que eu vou ficar mais para falar ali a questão do... Já ter, né, o que o algoritmo faria. Então, tá.

É, bom, o ponto de definição de requisito e, assim, das coisas que eu já faço, né, do meu trabalho atual, é, a gente sempre, assim, a partir ali do cenário, né, que precisa ser atendido, eu sempre vou a fundo no que realmente precisa, é, atender, né, a necessidade do cliente, no caso, né. Então, é, ele passou ali os dados, deixa eu pensar. Se quiser, eu posso voltar aqui no slide anterior para

você olhar quais são os requisitos.

Acho que eu vou precisar, mas eu tô pensando o lance do algoritmo, relacionar isso com o algoritmo. Tá, é porque também, supondo que seja, né, o primeiro anúncio para esse algoritmo que ser desenvolvido. Então, é, eu lembro até muito também da época da faculdade, né, de que ali a questão da análise de requisitos, de escrever, né, tudo que o sistema ou o algoritmo deve fazer.

Então, é, o algoritmo deve, é, publicar uma vaga, né, um post com os seguintes requisitos. E aí, o algoritmo deve ser capaz de permitir uma inscrição através de formulário. Eu acho que eu listaria, para aliar o levantamento de requisitos, né, eu listaria essas necessidades para, a partir daí, cada tópico, né, fazer uma construção.

Porque eu também imagino que não seria, vamos dizer, um script em si, vai ser um sistema, né? Então, acho que eu iria nessa linha. E, para a segunda pergunta, então, de dados, é, eu acho que faz mais, faz parte ali do requisito, né, que a gente definiria de início, né, se seria apenas uma publicação, enfim, um anúncio para captar pessoas se escrevendo. Um formulário, ou se é, ou se é a partir de um banco de dados interno mesmo, é, a gente, a minha empresa usaria, faria as filtragens dependendo, né, dessas especificações do anúncio.

Entrevistadora: Mas aí, os dados, por exemplo, específicos para alguma dessas características, né, que foram definidas, foram definidas em muito alto nível, né, mas você teria uma ideia, assim, de qual dado específico você filtraria para saber, por exemplo, formação, experiência, ou outras coisas nessa linha?

Leandra: É, acho que, não sei se você entendi, mas seria, por exemplo, então, nessa linha de, a gente já tem um banco de dados que as pessoas têm cadastro lá, né, então, os seus dados pessoais, e aí, sua experiência pessoal, seus skills, então, tendo, tendo esse, né, esse banco de dados, aí eu filtraria, faria ali o script, né, o algoritmo para filtrar. É, dentre as skills, né, quais, quais skills de, é, tanto, pois é, desenvolvedor é algo aberto, né, mas aí, dentro de full stack, estaria conhecimento de front-end, de back-end, banco de dados, é, então, o que eu vejo muito, por exemplo, é, é, a gente consegue colocar, em específico, né, por exemplo, pá, eu tenho conhecimento de linguagem Python, tenho SQL, tenho de JavaScript, né, são linguagens, mas, cada linguagem, a gente consegue associar, pois é, essa linguagem é mais usada para o front-end, essa linguagem é mais usada para o back-end, então, dependendo do que a pessoa listou ali como skill, aí, eu poderia fazer, dentro, né, uma parte lá do script, do algoritmo, um filtro para, para as skills que atendam, né, essa questão de ser full stack, por exemplo, a mesma, é, formação, formação seria, é, a mesma coisa, assim, dentre engenharia de software, aí, na lista, tem, na, nos requisitos aí, tem uma lista, né, engenharia de software, tecnologia de formação, sistemas ou afins, então,

eu penso que seria uma lista que, e de, e de formações que, é, tivessem, essa nomenclatura, né, engenharia de software, tecnologia de formação e os afins, pois é, pensando nesse, nesse, se for um sistema próprio, né, que, que o meu banco de dados já tem os valores, é, pré-estabelecidos, né, que, por exemplo, eu não consigo cadastrar coisa nova, o candidato, né, no caso, ele tem que escolher alguma coisa que está ali, então, já seria mais fácil de filtrar aquelas palavras-chave. Sim. Eu sou em algo nesse sentido, é, é, mas por ser afins, é, entra informática em geral, né, computação, é, fica realmente aberto.

Experiência, é, pois é, a, a partir desse ponto que as pessoas cadastram, é, como é que fala? O período, né, que elas trabalham, então, também seria um filtro em cima dessa parte, é, período, né, quanto, quanto tempo que elas têm ali cadastrado de experiência profissional. Então, seria algo, também, para refinar, né, com a empresa, com o cliente, né, que contratou, se, se ele quer cinco anos de experiência nessa, é, nesse tipo de, de cargo, né, de desenvolvedor, ou não, ah, cinco anos de experiência profissional, pode ser de, de outras coisas. Eu imagino, é que, assim, pensando na questão do óbvio, realmente parece, ah, cinco anos de experiência de desenvolvedor, mas é justamente o tipo de coisa que, aquilo que eu comentei, né, na hora de fazer o requisito, é, de ir no bem, bem especificar, né, detalhar e deixar claro, né, a necessidade do cliente.

Entrevistadora: Não, é, não é óbvio, né, é, tem muitas ramificações.

Leandra: É, é, é, não, é que, pois é, quando a gente lê, a gente já imagina, né, se eu vou lendo desde o início, ah, fica meio que óbvio que é cinco anos de experiência em tecnologia, mas, mas não, não necessariamente, né, as pessoas, elas podem vir, pois é, de administração, enfim, de outro ramo, tem só um ano de experiência de desenvolvedor. Enfim, é, e regime remoto, é, também seria um, um, pensando nesse contexto do, de até as informações ali, né, de até o banco de dados com que as pessoas cadastraram, aí teria uma flag para os candidatos colocarem, se, se estão abertas a, a remoto, ou não e, enfim, também faria isso, que seja o filtro nessa, para essas informações.

Entrevistadora: Legal. Daí assim, pensando já mais no desenvolvimento do algoritmo mesmo, você consegue me dizer algum método que você escolheria para o algoritmo específico, para montar o algoritmo? E se tem um desafio técnico que você já vem falando, né? Mas tem algum que norteia, algum que pode nortear a escolha do método, né? Dependendo do desafio técnico, você já vai mudar a escolha do método de princípio.

Leandra: É, tá. Eu acho que eu já fui levando o meu pensamento, mas vou dizer simples, né? Seria mais algo de filtragem, de ter um banco. Eu acho que é porque eu já estou vendo ali o negócio técnica de machine learning, e assim, eu não tenho tanta experiência, né? Com essa parte, mas deixa eu pensar.

Não, mas não precisa mesmo. É assim, dentro do seu conhecimento, a gente sabe que tem várias possibilidades, né? E que você acha que, vindo à cabeça, a princípio assim, poderia contemplar essa questão técnica?

Leandra: Eu não sei dizer. Eu tô pensando muito, acho que muito mais.

Eu não sei se tem nome essa técnica, não lembro. De a partir de um banco de dados, né? Então as pessoas se cadastrando antes, você trabalha ali com filtros, né? E tudo mais. Mas eu realmente não pensei a questão de usar, então, pois é, alguma ML para fazer essa seleção, né? Não, não tem problema.

Não precisa também ser machine learning, mas... É porque, pois é, uma machine learning, a gente conseguiria, como é que eu vou dizer? Desenvolver ela para, como é que fala? Para ela selecionar aqueles requisitos, né? E aí ela já trazer a partir do banco de dados, né? Aquela, os candidatos que se encaixam naquele perfil. É, realmente, eu acho que eu tô com muita mente de botar os dados disponíveis, né? Mas a partir dali, realmente, eu não tive essa experiência de projeto.

Entrevistadora: Tranquilo.

E aí o resto também, assim, não sei se você pensa em alguma maneira de testar para validar, né? Esse modelo, se tem alguma métrica que você acha que vale usar, aplicar. Até pensando naquela situação inicial, né? Você foi, sua empresa foi contratada. Então, tem que mostrar que aquele algoritmo está funcionando, dando resultado, né? Quais são os resultados.

E se você considera ter feedback para essas interações, né? Pode ser humano, não humano, tanto faz.

Leandra: Sim, sim. Ah, eu diria, deixa eu pensar.

Eu vou dizer algo bem básico e eu tive algumas experiências de fazer projetinhos pessoais mesmo, com o IA, com machine learning, né? E tudo mais. E tinha bastante essa questão do teste, né? Vou dizer que verificar ali o resultado, né? Vou dizer manualmente, assim, pegar uma amostra e verificar se aqueles perfis que apareceram no final correspondem, né? Aos filtros, né? Que a gente tinha colocado antes. E também mostrar alguma amostra para o cliente, né? Para ele até estar aqui.

Isso também seria a parte de um feedback, né? Verificar se é aquilo que está atendendo. É, mas aí os testes mais específicos, né? Que eu sei que existem. Teria que pesquisar para poder, de fato, implementar.

Métricas de sucesso. É, também isso teria que ir mais a fundo.

Entrevistadora: Meio junto, né? Vai meio junto ali.

E feedback? Você acha que na hora de montar, tomar as decisões, pensar em quais dados você pegaria?

Leandra: Ah, eu acho que tem que ser constante com o cliente e também entre o time, né? Porque as pessoas... Eu imagino, né? Que eu vou estar desenvolvendo junto, né? Com outras pessoas. Inclusive de pessoas que foram selecionadas, né?

Entrevistadora: Legal. Tá.

E aí, assim, pensando já em uma coisa até mais, vamos dizer, abrangente, né? No nível sociedade, assim. Você acha que o algoritmo, dependendo do jeito que você modelar, do jeito que você configurar, ele pode gerar um impacto na contratação positivo, negativo? Como seria, né?

Leandra: Ah, sim. Acredito que sim.

Pode haver impacto. E me veio na cabeça agora justamente a questão de ter que ter cuidado, né? Na hora, com os vieses, né? Na hora de fazer o algoritmo. Pra, por exemplo, ele não ir pra um lado de que, ah, são essas skills aqui, tem mais homens que realizam.

Então, vou selecionar mais homens porque os homens, então, que são qualificados. Teria que ser algo, pra estar atento, né? Se o algoritmo não tá, não tá com esse viese, né? Enfim, não é porque tem muitos homens que se inscreveram e que colocaram ali que eles têm essa skill que eles sejam melhores, né? Uma coisa diferente.

Entrevistadora: E você considera que você, como desenvolvedora, você deve alertar para possíveis vieses, trazer outras pessoas para tentar melhorar, né? Essa visão?

Leandra: Com certeza, com certeza.

Eu acho, acredito muito naquilo de ter uma equipe multidisciplinar. Multidisciplinar, é, mas até a questão cultural, né?

Entrevistadora: Diversa.

Leandra: Diversa, isso.

Obrigada. Pra, justamente, perceber, né? Porque, às vezes, coisa mesmo que seja, ah, são mulheres, tem alguma coisa ali que elas podem não perceber. Então, acredito que faz diferença, né? Na hora do, na hora de desenvolver e mais do que desenvolver, né? Na hora de ver o resultado.

Entrevistadora: Sim. E você acha que se tivesse um framework para te direcionar, pensar em algumas questões, isso te facilitaria?

Leandra: Sim, sim. Porque, é, o framework, ele vai, vai ser tipo um checklist, né? Assim, ele vai, faça isso, enfim.

Sim, acredito que auxiliaria nessa construção.

Entrevistadora: Legal. Leandro, muito, muito obrigada.

Aqui acaba, é, foi bem legal ver o seu pensamento, você se preocupando realmente muito com uma questão de funcionar mesmo, assim, né? Uma

questão técnica e é cheio de desafios mesmo essa solução, não é nada simples, né? Não é nada simples. E foi legal esse ponto que, esses pontos mais de viés, né? Que você trouxe, não é todo mundo que pensa, né? A princípio, porque nem sempre é a realidade, né? De cada um. E acaba que a gente traz bastante da nossa realidade também, né? Eu vou passar aqui o link do formulário, tá? Quando você puder responder, não sei se é agora, não.

É bom. É rapidinho, mas é só para complementar, né? Os dados. E deixa eu te perguntar, você mora no Rio?

Leandra: Aham.

Entrevistadora: Você vai quarta-feira?

Leandra: Não sei, tem coisa quarta? Ou você acostuma a ir quarta-feira, né?

Entrevistadora: É, porque meio que assim, as pessoas, pelo TIM, tinha o sócio, né? Da nossa frente aqui, ele meio que fala assim, ah, quem puder ir na quarta, assim, o gerente, ele meio que obriga. Mentira, não obriga, mas ele gosta de ver o pessoal lá. Mas a galera do staff, ele fala, pô, quem quiser vir e tal, é legal e se não for muito longe para você, difícil de chegar, é sempre, é legal a gente socializa, né?

Leandra: Não, eu quero muito, eu só não consegui até a Ju, né? Falou para eu ir e eu tinha falado para ela que eu ia, né? Na semana que eu entrei, só que foi muito corrido, né? Eu fui na segunda, né? Para pegar a máquina e aí chegou quarta-feira, eu falei, Ju, não vai dar para eu ir, vou deixar para semana que vem.

Só que a semana passada eu também não consegui ir, mas aí essa agora eu já conseguiria ir, então já vou estar lá, aviso para ela também e eu acho que o TIM chegou aí na semana passada?

Entrevistadora: Não, semana passada não, na outra que ele foi, na anterior.

Leandra: Mas eu só fiquei com isso na cabeça porque eu acho que teve algum evento e aí ele estava no aeroporto. Eu falei, caraca, ele não deve ter conseguido ir para o escritório.

Entrevistadora: Eu acho que ele foi para a Califórnia, algum evento da NVIDIA, alguma coisa assim.

Leandra: Ele comentou, é verdade. Mas aí, enfim, também não deu para eu ir, mas aí essa quarta eu vou lá então, a gente se vê pessoalmente.

Deixa eu te perguntar, mas aí você vai toda a quarta?

Entrevistadora: É, não é toda não. Eu não me boto uma pressão de ir toda a quarta não, mas para mim é muito fácil porque eu moro perto, sabe? Moro ali na Glória, muito pertinho.

Leandra: Eu sou de Madureira.

Ah, legal. Não é tão perto, mas eu moro perto da estação de trem, então para mim eu pego o trem. E aí lá eu vi que também dava para pegar o VLT, dava até para ir andando, né? Aqui no primeiro dia eu acabei pegando o VLT, mas é tranquilo.

Entrevistadora: É melhor, né? Se bem que da central para lá de dia é de boa, o negócio é mais para voltar. É, é. Mas a gente sai junto ali também. Ah, legal.

Mas é legal, assim, eu não vou toda a terça, a quarta não. Tem dia que eu não dá. Às vezes, o que acontece quando a gente vai, a gente não produz tanto, sabe? Ah, é. Aí se eu estiver muito apertada, mas assim, as conversas que saem, a conexão que a gente faz é bem legal, assim.

Eu gosto de ir.

Leandra: Eu estou doida para ir. E deixa eu te perguntar, porque também como eu sou nova aqui, né, e vendo as coisas agora da Deloitte, eu estava tentando pensar se valia a pena pedir, por exemplo, o Valetransporte.

Porque por mais se botasse uma vez na semana, né, que seria quatro vezes no mês...

Entrevistadora: Eu acho que não. Sabe por quê? Valetransporte diz conta de uma porcentagem de seu salário. É, né? Eu não sei como é que é hoje, tem muito tempo que eu já vejo que não vale, sabe? Porque, sei lá, sete por cento e aí fica só para Valetransporte, sabe? E aí seu valor que vai pagar, né? E aí você... É, mas eu estava a perguntar para alguém que fosse e se... E que achava, porque realmente tem muito tempo que eu não... Eu não entendo o Valetransporte, nem lembro.

Valetransporte realmente não é uma coisa tipo extra, sabe? É uma coisa que desconta do seu salário. E aí, a não ser que você vá todo dia, que seja uma passagem, sei lá, intermunicipal, muito mais cara e tal. Mas eu acho que assim, uma vez por semana, sem obrigação, porque aí tem semana que você não vai querer ir e tal.

Acho que não vale, não. É, né? Ah, tá bem. Tranquilo.

E aí depois, se você for a gente conversa mais também, mas queria te pedir mais uma coisa. Se você tiver colegas, mulheres, né? Que têm o seu perfil, né? Que desenvolvem algoritmos e tal. Se você conseguir me passar, me ajudar a conseguir mais gente, eu agradeço.

Sim, deixa eu te perguntar. Mas aí, você está focada no pessoal aqui da Deloitte? Não. Ah, interno também? Também.

Ah, então. Não, tenho demais, porque... Eu não falei aqui, né? Mas eu participo de comunidades... Inclusive, eu conheci a Ju assim, né? Não sei se

ela falou. Ah, que legal.

Foi numa comunidade de mulheres de 100 anos de idade. Então, aí tem várias pessoas lá, né? E mais próximas, né? Para indicar. Ah, nossa.

Eu vou fazer o seguinte. Eu vou montar uma chamadinha, aí eu te faço. Se você puder compartilhar, vai ser maravilhoso.

Ótimo. Porque aí a pessoa já bota o nome, né? E contato para você. Ah, tá ótimo.

Fechou. Obrigada, viu? Nada. Tchau, tchau.

Tchau, tchau. Boa tarde. Boa tarde.

B.16

Luciana

Entrevistadora: Tá, acho que tá, tá tudo certo. Bom, eu vou compartilhar a tela com você, mas a ideia é o seguinte, são apenas direcionadores, assim, para a gente ir pautando a nossa conversa, tá? Eu vou usar a metodologia de, o método de entrevista semi-estruturada, então tem, não é pergunta e resposta, tá tudo bem aberto, e conduzindo também com o método de direct storytelling, que é um método do design, que a gente pede para a pessoa realizar uma ação, e aí é importante que você vá falando tudo que tiver a cabeça, assim, então, realmente eu vou analisar muito mais as questões do, de, esse modelo mental que você vai construir para resolver essa, trazer essa solução, do que de fato as questões técnicas, tá? Então, se você quiser compartilhar a tela, para mostrar alguma ferramenta, qualquer coisa, fica à vontade, pode pegar aqui a hora que você quiser. Então, é, eu vou pedir primeiro para você fazer uma contextualização da sua experiência, né, em relação a, com algoritmos da experiência que você já teve, sua experiência profissional, pode ser acadêmica também, e dizer um pouco do seu papel na empresa hoje que você tá, mas assim, não precisa falar de projetos, nada muito aprofundado, só de uma maneira mais ampla mesmo.

Luciana: Então, eu tenho uma formação de estatística, o que faz eu, né, conhecer de algoritmos, e eu trabalhei a vida inteira com análise de dados. Boa parte do meu trabalho, como a maior parte do trabalho de analistas é muito mais, tá muito mais focada em análise descritiva, do que análise diferencial, mas eu tive várias oportunidades para aplicar análises diferenciais, seja para previsão, ou para segmentação, que muito mais como classificação, acho que eu tenho muito mais experiência em situações diferentes, seja em pesquisas de dados de pesquisas de mercado, seja dados de clientes, seja criar clusters de negócio. Então, isso, tenho aí os últimos 15 anos, fiz isso colocando a mão na massa, ou gerenciando e criticando os algoritmos do time que faz.

Hoje, eu gerencio um time de analíticas, que tem 16 pessoas, entre analistas e engenheiras de dados, um dividido em três equipes, e uma, é, três lideranças, não exatamente equipes. Uma dela é muito focada em engenharia de dados mesmo, então, em gestão, disponibilidade de modelo de dados, mais modelo, né, de acesso ali, montar o ETL e disponibilizar a informação, mas as outras duas lideranças ali são de times mais analíticos, de fato. E aí, as nossas, aqui, especificamente no meu time, a gente tá muito direcionado, ou para marketing, ou para o que a gente fala, que é o topo do funil.

Então, aqui no Quinto Andar, o topo do funil é quando, ele tem dois grandes públicos, que é o proprietário ou o inquilino comprador, né. O proprietário, o topo do funil dele, é ele publicar o imóvel com a gente, então, conseguir trazer essas pessoas para publicar, e eles conseguem publicar, de fato. E o inquilino ou comprador, é ele chegar aqui, entrar no site, pesquisar, até ele agendar uma visita.

Aí, o fim do funil dos dois é o contrato assinado, né. E aí, os funis se encontram e o contrato se é assinado, aí já vai, outras áreas olham para essa parte. Eu foco muito aqui em encher o baldinho, né.

Entrevistadora: Entendi. Luciana: Vai encher o balde.

Entrevistadora: Muito bom.

Tem a ver um pouco com esse desafio que eu vou falar agora. Eu vou falar um cenário fictício, tá, e aí vou te explicando, à medida que você tiver dúvida, pode perguntar. Imagina que uma empresa de tecnologia, de porte médio, tá com uma vaga aberta para contratar, né, um profissional para a sua, para o seu corpo ali de profissionais, e contrata uma outra empresa para fazer a divulgação dessas vagas, tá, dessa vaga específica.

Você vai ser responsável por fazer esse anúncio rodar e chegar nas pessoas certas. O que seriam as pessoas certas? Aí você vai modelar, né, vai fazer um modelo para minerar essas pessoas que tenham perfis que dêem fit com essa vaga que está sendo anunciada. Então eles passam aqui esses requisitos bem alto nível, uma vaga de tecnologia para profissional pleno, com skills para desenvolvedora ou desenvolvedora full stack, com formação em engenharia de software, tecnologia da informação, sistemas e afins, uma experiência de cinco anos, e para trabalhar regime remoto.

Então eles passam esses requisitos, contratam você e fala, faz esse anúncio chegar nas pessoas que podem ter interesse e que, uma vez entrando nesse funil, que lá na frente seja convertido em contratação. Então aqui, já pensando nessa simulação para essa solução, no princípio aqui de planejamento e concepção, né, como você definiria esses requisitos, né, como você aprofunda para refinar ali os requisitos, e uma vez com esses requisitos refinados, quais dados você

utilizaria? Quais dados você acha que podem responder às questões que você vai colocar? Vou voltar lá no primeiro para você olhar os requisitos.

Luciana: Uma coisa que, assim, é um desafio de topo de funil, é que a gente tem menos informação, normalmente.

Então, por exemplo, eu tenho muito, pensando aqui no caso do Quinto Andar, eu tenho muito mais informação de quem está fazendo uma aprovação de crédito, né, ele já deu muita informação para a gente, do que uma pessoa que simplesmente entrou no nosso site e saiu. Pessoa que entrou no meu site e saiu, eu tenho um cookie com alguma coisa de informação dele, sei por onde ele navegou e olha lá. Quando a pessoa se candidata para uma vaga, eu não tenho, eu tenho só o currículo, o que ela pôs no currículo dela, né.

Ela vai falar, dá cheque aqui e não dá. Então, mas eu não tenho o que ela me trouxe em entrevista, eu não tenho, eu não sei se você está contando aqui que já nesse processo você faz algumas, é só cadastro de currículo mesmo, né.

Entrevistadora: É, a pessoa aparecer para ir para o RH lá avaliar.

Luciana: Para fazer um filtro para o próximo passo, né.

Entrevistadora: Isso.

Luciana: É, então, aí para o próximo passo já vai ter mais informação.

Se você tem mais informação, você vai, quanto mais informação você tem, né, estatisticamente, você tem mais probabilidade de ser mais assertivo, né. De pegar as variáveis mais importantes. Se você tem mais informação, você pode olhar, por exemplo, a... Não só contratar, mas fazer uma correlação com que perfis performam bem.

Pode ter mais. Aqui você vai olhar, basicamente, corresponde ou não corresponde aos requisitos. E se eu tivesse vários requisitos, quais deles funcionariam? Quais deles a gente tem mais sentido contratar? Então, as variáveis, é fazer uma mineração nos currículos inscritos para tentar entender quais estão, quais tem fit com a vaga de fato.

Até porque, putz, tem uma parte que é muito algoritmo... Nossa, falhou o nome. De recomendação? Não, que é quando... Você faz valorística, você está fazendo regras muito claras. Então, cara, tem de fato cinco anos de experiência? Chegou aqui com um ano, cortou.

Está tirando, realmente não fez requisito. Você pode criar ali uma regra para ir filtrando isso. Só que aí vai sobrar um monte de gente que corresponde a tudo isso.

Como que você faz para separar para a próxima fase de RH? Necessariamente eu vou falar que mais anos de experiência é melhor? Não necessariamente a gente vai fazer isso. Então, olhar para quem a gente já contratou com essa vaga, porque geralmente quando a gente faz um algoritmo supervisionado, você

coloca alguns casos de sucesso como referência. Quem você contratou, quais eram as informações que batiam? Então, por exemplo, a formação... A gente vai ter várias formações diferentes.

Aí você vai filtrar, colocando que o algoritmo vai dar uma pontuação maior para quem tem algumas formações. Só que aí você começa a entrar em alguns problemas também nisso. Um exemplo.

Lá atrás, na minha carreira, era muito comum os filtros de empresas pegar assim. Formou na USP, na Unicamp, no FRJ e o resto cortava fora. Pegava algumas pessoas assim.

E aí você está pegando um perfil de gente, né? Claramente você está deixando um público de fora. E quem disse que esse público de fora não é tão bom quanto? Então, o ponto é, eu posso fazer... Ah, eu tenho muita gente aqui boa que tem essa formação de escola de elite aqui dentro, mas o quanto outras pessoas de outras escolas não entraram. Então, mesmo que eu faça ali um algoritmo em caixa preta, eu só coloque a saída... Coloco as minhas casas de sucesso e a saída sem eu conseguir ver qual é o peso das variáveis, eu tenho que fazer depois uma análise dos perfis para ver que perfis estão sendo selecionados para entender quais variáveis estão pesando.

Então, tem algumas técnicas. Fazer um shape value por cada variável que entrou para entender quais que pesaram mais no algoritmo de classificação. E ter crítica em cima disso.

Não é porque o algoritmo soltou que a resposta do algoritmo é boa, né? Então, pô, está me dando um peso enorme na escola... Na faculdade, por exemplo, esse exemplo que deu de qual a universidade que se formou, porque essa é uma informação que eu tenho aí no topo do funil, né? Que a pessoa colocou no cadastro do currículo. Eu quero que isso esteja na minha informação. Pô, isso está diferenciando um pouco.

Talvez faça sentido. Mas isso está diferenciando muito mais que todo o resto. Toda vez que uma variável tem peso no modelo, eu tenho que olhar para essa variável com crítica dela.

Então, é isso. Eu tenho muito pouca informação. Eu tenho muito pouca informação, mas eu não posso deixar de ser crítico, né? O que eu colocaria é... A minha base de treinamento tem informações de quem entrou aqui, de processos que eu considero sucesso, né? Mas a minha base de aplicação, eu tenho que olhar quais são as variáveis que estão diferenciando a resposta desse algoritmo.

Entrevistadora: Entendi. Legal. E aí, nesse caso, quando você fala em fazer essa análise, quando você percebe que é uma coisa que impacta, né? O peso ali que você colocou.

Você costuma discutir entre pares, com as pessoas do grupo? No caso, você é gerente, você discute com eles? Tem essa rotina, né? Como é que é?

Luciana: Tá. Uma coisa que eu sempre cobro muito do meu time é que a análise não pode sair da análise de dados. A gente tem que saber mais o negócio para conseguir discutir.

Então, não é pegar a task e pegar pronta, né? Eu tenho que fazer uma segmentação. Por exemplo, seja essa aqui a task. E no final, eu devolvo.

O algoritmo é esse. Dá o play aqui, sei lá. Coloca essa base.

Eu vou te dar um notebook no Databricks que você roda e vai cuspir quem são as pessoas dela. Aí você tem que discutir com a pessoa, porque quem vai entender mais de seleção? Eu espero, né? Que a pessoa de RH entenda mais de seleção do que eu. Por mais que, vamos supor, eu atendesse People Analytics, que é uma área de expansão dentro de Analytics.

Eu entendo. A pessoa de RH entende mais. Então, durante o processo, durante filtro de variáveis, durante toda a análise, na discussão disso, você tem que estar junto com quem entende daquilo.

Então, aqui, por exemplo, no meu caso, eu faço muito isso para definir formatos de mídia digital. Cara, eu não posso simplesmente com a análise de dados decidir os formatos de mídia digital. Tem que ter uma análise de mídia ali na discussão.

Ele não vai rodar o algoritmo, mas ele tem que conseguir entender o básico para fazer a crítica das variáveis que soltam. Porque, às vezes, eu tenho um modelo com uma assertividade, e aí a gente pode falar que é uma acurácia estatística melhor, né? Mas não necessariamente ele funciona na vida real. É que nem, sei lá, um teste de laboratório de... Que, em vitro, funciona, mas quando faz o teste de remédios, faz o teste em vitro, funciona muito bem.

Por isso que é obrigatório. Não pode fazer só em vitro. Tem que testar em pessoas, de fato.

Aí tem um monte de coisa que funciona em vitro e não funciona em pessoas. E o algoritmo é isso. Muitas vezes vai ter uma acurácia estatística, mas, na prática, vai trazer algumas coisas que não têm sentido.

Então, você tem que ter uma crítica língua. O melhor algoritmo não é o da melhor acurácia. É o que vai dar... É o que tem o melhor resultado na vida real.

Entrevistadora: Tá, entendi.

Luciana: E, assim, em vida real, eu teria que testar vários. Então, eu tenho que misturar as informações estatísticas com esse feeling de negócio e entendimento.

E quanto mais gente eu tiver ali discutindo, aí é o trade-off entre ter muita gente discutindo ou ficar numa eterna discussão também. Uma hora tem que ter um bom final. E é isso, gente. Vamos partir daqui.

Entrevistadora: Entendi. Então, nessa parte, você falou já alguma coisa, né? Mas você conseguiria dizer... Consegue falar algum método que você escolheria para fazer esse algoritmo? E se algum desafio técnico que você já consiga antever, faria, até direcionaria a escolha do seu método, né? E essa integração, implementação, no caso, assim... A gente está falando, de repente, de aparecer um anúncio no LinkedIn, Facebook, sei lá, né? Se você tem uma ideia de como seria.

Luciana: Quais são os métodos técnicos de machine learning? Bom, aí daria para usar, sei lá, uma regra só logística. Que é um método de classificação bem... Que costuma se adaptar bem nesse tipo de cenário. Uma árvore de decisão.

Ou uma rede neural. Mas, geralmente, até como hoje, está cada vez mais fácil aplicar esses métodos. Geralmente, a gente faz vários.

E vê qual funciona melhor, mais do que escolher o método a priori, né? A gente faz um conjunto de métodos. A definição antes é, eu quero fazer previsão, ou quero fazer classificação. Aqui, a classificação seria, próximo passo ou não, né? Ou um score de... Ordenar quais são os melhores casos ali.

E aí, por exemplo, uma regressão logística poderia me dar esse score. Desafios... Mas aí, roda vários algoritmos e vê qual funciona melhor. Desafios técnicos durante o desenvolvimento do algoritmo.

Bom, a gente está falando de dados que são não estruturados. Então, tem uma parte aí técnica de como trabalhar esses dados. Porque a gente está falando de dados que não são dados de texto.

Então, pode ser que essa aí... Porque toda vez que a gente fala, trabalha com dados não estruturados, no fundo, você está pegando eles e estruturando. Tem um algoritmo de estruturação e depois sai. E muitas vezes, falha aí nessa estruturação.

Por exemplo, quando você pega currículo, essas pessoas, as pessoas cada vez estão mais criativas no formato do currículo. Quem disse que as plataformas estão preparadas para ler isso? Entendi. Então, muitas vezes, assim, eu estou pegando ali como uma informação importante, eu vejo que a pessoa ter colocado um skill de métodos ágeis dentro do currículo dela é uma variável que está dando diferença.

Mas se as pessoas colocaram isso de um jeito que... Se alguns currículos colocaram isso de um jeito que a plataforma não lê, não consegue estruturar que aquela pessoa está classificada como método ágil. Então, a gente tem que

ter um trabalho de auditoria grande nesse processo para entender se... O desafio técnico, o primeiro que eu vejo é garantir que a técnica está funcionando na leitura. Muitas vezes, a gente vê o problema lá no final do algoritmo e esquece esse começo.

A estruturação. Então, eu consigo ler os dados devidos ali da forma certa. Depois, eu acho que é um desafio técnico.

Eu falei, quando a gente está no topo do funil, geralmente a gente tem muito menos informação do que a gente gostaria para fazer. E você falou que é uma empresa de médio e forte. Eu estou falando que eu preciso de uma base de treino que tenha umas entradas ali.

Eu não sei qual é o tamanho da amostra que eu já tenho na empresa. Então, se foi uma empresa de médio e forte, sei lá, se é uma startup que bombou em dois anos, será que eu tenho base suficiente para falar que eu tenho dados para treinar um modelo? Então, pode ser que eu não tenha dados suficientes. Uma empresa mais consolidada, com uma área de tecnologia grande, provavelmente eu tenho.

Ou, se for o caso, talvez pegar dados de outras frentes de mercado. E a integração e implementação. Se, por exemplo, divulgou no LinkedIn, alguma coisa.

A maioria das empresas hoje usa aquele sistema, sei lá, o Gupy, ele, o Workable, alguma outra disso. Provavelmente você vai fazer a leitura da API dessas ferramentas. Então, como eu falei, primeiro essa ferramenta tem que conseguir ler certo.

Você faz a leitura a partir dessas, da API dessas ferramentas, traz para o seu ambiente. Provavelmente, então, qual é o ambiente que você consegue trabalhar? Você vai trabalhar, por exemplo, vou conseguir ler isso, jogando esses dados no meu Data Lake, lendo pelo, fazendo um algoritmo em um ambiente da AWS. Vou fazer algo assim.

Acho que a integração é essa coisa mais sistemática. Depende da arquitetura, da stack de dados que a empresa tem, vai depender a implementação, a integração. Mas o importante da implementação é, acho que o desafio de implementação é fazer essa auditoria dos passos.

Porque eu vejo, muitas vezes, o negócio ir lá no final, e acho que esse exemplo que eu dei é muito bom, de, pô, lá no final, a classificação não está tão boa. Mas eu vou ver, desde o começo eu não estou conseguindo ler que aquela pessoa tem um skill de método ágil. Então, vou ter que, na implementação, garantir que cada passo da implementação está tendo um output devido.

Entrevistadora: Entendi, bom. E aí, em relação aos testes e validação, se você tem alguma ideia de teste que você poderia realizar? Algumas métricas de

sucesso, lembrando que é um produto que você está vendendo para uma outra empresa. E se você considera feedbacks para interações humanos, internos, externos?

Luciana: Melhor dos mundos é quando a gente consegue, para qualquer teste, melhor dos mundos é quando a gente consegue isolar um grupo teste controle, separado um teste A, B, de fazer, seja, rodar algumas vagas com e algumas vagas sem esse algoritmo, e aí definir quais são as métricas de sucesso.

Ah, qual que é o esforço, sei lá, uma métrica que a gente coloca esforço da recrutadora, horas de recrutadora por contratação de sucesso. Quantas horas destinadas ali para isso, como que fazia no método atual, como que faz no novo, e roda isso por uns meses e avalia o resultado e coloca algumas métricas de sucesso, por exemplo, isso. Ah, essas pessoas foram razoavelmente bem habilitadas no tempo de experiência, os 90 dias.

Algumas métricas de Algumas self metrics que a gente coloca ali para ver que não são realmente exatamente o sucesso, que a gente vai falar do teste dados significantes ou não, mas que a gente tem que avaliar se está ok. Então, sei lá, eu vou Aí muito tem que discutir um pouco com o pessoal de Arragal, o que seria ali Algumas self metrics, será que eu estou trazendo só pessoas que estão com um salário muito acima do mercado, será que eu estou trazendo pessoas só de uma mesma empresa? A gente não consegue porque quando a gente traz gente de fora, a gente espera também que tem um refresh ali, né? Ou será que eu estou trazendo com isso só um perfil, estou excluindo alguma questão de diversidade ali no meio? Acho que teria que olhar isso, mas muitas vezes a gente não consegue implementar um teste, assim, o mundo ideal é fazer sempre um teste aberto, caso controle em tudo, mas não necessariamente também eu vou conseguir fazer isso, porque eu teria que ter poder ser estatisticamente significativo eu teria que contratar muita gente ter um número de vagas grandes, então talvez eu não conseguisse fazer isso nesse mundo ideal, né? E aí, algumas vezes o que a gente faz é um o que a gente já fala que é o pseudo teste que a gente não está fazendo mais aberto, a gente faz como que funcionou antes e depois, faz o antes e depois Só que depois eu não estou mais isolando, eu tenho a influência da variável tempo, pode ser simplesmente que chegou mais gente boa

Entrevistadora: Entendi

Luciana: Aí não tem nada a ver com o meu algoritmo se eu separo, aleatoriamente está dividido, eu consigo tirar o efeito dessa variável dessa cor variável, mas eu não consigo tirar, mas às vezes é o que a gente tem a fazer, e vê se tem e pelo menos essa visão de hora executadora, se está aí quando a gente não faz, quando a gente faz um pseudo teste, a self matrix ganha mais

importância normalmente, né? A gente consegue ter uma significância de fato ali Acabou que eu falei ali um pouco de métrica de sucesso, tendo que é tempo de porque contratar em tecnologia uma das dificuldades é muito tempo, muita busca então acho que talvez essa esse tempo de contratação de sucesso poderia ser uma métrica, e depende dos direcionais da empresa a empresa está querendo trazer pessoas com que tenham mais por mais que sejam técnicas, tem uma postura tem um perfil mais business um perfil mais de discussão estratégica é uma coisa que você não vai conseguir avaliar nesse currículo, geralmente você vai avaliar em entrevista e tal, mas você consegue avaliar um pouco em uma primeira avaliação da pessoa, será que você está conseguindo, será que esse método não está filtrando isso independentemente de alguma forma, ou será que esse método é só homem em trazer mulher será que esse método está cortando alguma outra questão de diversidade que, nossa de repente não entra mais negro aqui na empresa, porque todos os negros foram filtrados nisso filtrados por causa de uma questão que mais tempo de experiência, tem menos gente, foi filtrado pelas universidades que fez, por exemplo tá lá que as universidades cortam, é uma coisa um pouco mais recente tá perdão, pode falar pode levar a modificações do algoritmo essas interações esses pontos que eu falei de ir trocando com a área de negócio que é dona desse produto quem é o usuário, o demandante do produto, tem que ser feito durante, desde o planejamento até a avaliação do algoritmo e durante a implementação do resultado teste, não é um não é um produto de prateleira toma aqui a área de dados faz e devolve, tem que ser construído a quatro mãos.

Entrevistadora: E aí por fim em relação aos resultados e impactos, mais uma opinião mais pessoal sua em relação a efetividade como você avalia que essa a efetividade do algoritmo não necessariamente de métricas, de testes ali feedbacks, mas você acha que isso é efetivo no contexto de contratação fazer a utilização desse algoritmo para essa primeira filtragem e se isso impacta na contratação de uma maneira ampla algum impacto significativo social cultural, questões que você também já trouxe que podem impactar

Luciana: É efetividade eu como analista de dados sempre parto do não sei eu tenho que estar na hipótese nula, a gente tem tem que comprovar a hipótese eu tenho que analisar ela então não sei se isso funcionaria eu tenho que testar e aí é isso Olha o que o mercado já fez disso já vi exemplos de mercado disso funcionar bem, já vi exemplos disso não funcionar bem tem um case que há muitos anos eu uso dando aula, que é um case da Amazon que a Amazon fez isso, um algoritmo de contratação quando eu falo sobre a importância do que entra de variável que as vezes a gente confia muito no output do modelo e esquece que tem um input ali dentro e do histórico da base de pessoas, tem um

case da Amazon de, sei lá, 10 anos atrás que eu uso em aula muitas vezes que a Amazon fez um algoritmo de contratação e eu acho que eles usavam até como variável resposta pessoas que performavam bem na empresa e tal e no fim das contas nenhuma mulher era selecionada nesse algoritmo Ah, é o algoritmo que é machista? Não é o algoritmo que é machista A história da empresa é machista e tem muito mais homens ali dentro então o algoritmo não vai entender que quem está performando bem é homem, é isso o algoritmo está performando bem, só que no fim das contas ele pegou a variável homem e colocou ali então a gente tem que ter essa crítica, esse caso é bem antigo e bem clássico assim, quando a gente fala de crítica quando a gente fala de crítica em algoritmo nem falando exatamente desse exemplo de contratação mas nesse caso até bate, mas esse é um exemplo bem antigo e bem difundido mas ao mesmo tempo a gente tem a gente tem uma impossibilidade é muito difícil a contratação, é um esforço muito grande, a gente sabe que não dá pra a gente tem que ter técnicas para triar, fazer triagem de currículos, aqui por exemplo a gente está com várias, é sempre uma briga quando tem muitas vagas abertas ao mesmo tempo porque não dá e chega muita coisa que não tem sentido para a contratação, então pode sim fazer, eu acho que testar ver as referências de mercado, tem casos de sucesso e casos de fracasso com isso isso deveria dar um, que assim a gente deveria ter um ganho de menos esforço para conseguir trazer alguém bom se eu não tiver menos esforço aí a tecnologia não vai trabalhar no meu favor eu tenho que conseguir acho que o impacto é, eu consigo ter menos esforço para trazer pessoas boas mas o ponto que eu penso é que muitas vezes a gente pensa no a gente pensa no endgame, o endgame é ah, eu vou pegar currículos, eu vou falar quem que eu deveria contratar e às vezes sempre tem um meio do caminho que faz muito, às vezes faz muito mais diferença do que coisa, por exemplo ah, chegou a 300 currículos aqui e eu preciso mandar 20 para a entrevista será que eu tenho esse algoritmo, eu posso fazer um algoritmo que só pega desses 300 currículos e filtra os 100 melhores, então em vez de mandar 20, manda 100 e aí tem uma visão humana em cima desses 100 será que esse não seria um caso eu tenho um pouco mais de esforço mas eu tenho uma redução de horas bem considerável entre um e outro, porque afinal eu não estou lá mais lendo 300 currículos, eu estou lendo 100 e talvez o resultado final na soma seja bom às vezes a gente olha muito isso, sabe o melhor cenário, a maior automação de todas pode fazer uma automação intermediária com inteligência para fazer uma limpeza básica como por exemplo, eu falei lá, um dos requisitos era 5 anos de experiência será que precisa alguém ler para saber que tem 5 anos de experiência é muito fácil de implementar no algoritmo você já consegue filtrar muita coisa e eu acho que muitas vezes acontece se esquece isso nesse boom

de usar muita inteligência artificial para tudo, se esquece que às vezes você pode usar um pouquinho de inteligência e no fim o ganho pode ser maior não exatamente na variável que você está pensando mas no fim, você não cai nesse monte de erro do algoritmo que talvez você não tenha base para isso e você tem um ganho final, você tem um filtro você tem um ganho de eficiência lá no seu time de recrutamento.

Entrevistadora: Uhum. E uma pergunta final, você considera que seria interessante ter um checklist já que a gente está falando de algoritmo que vai impactar em questões sociais um desenho ali de pessoas que vão ter uma oportunidade ou não e aí tudo isso que você já trouxe também a questão de acesso e tudo mais você acha que seria importante nessas situações em que o algoritmo tem uma agência grande social, ter um checklist com questões direcionadores de diversidade, inclusão, questões que deixem, que façam o norteamento de um produto, um desenvolvimento de algo mais inclusivo

Luciana: Acho que é bom porque para todo mundo eu espero que esteja todo mundo evoluindo com isso. É algo que as pessoas como eu falei, é muito comum ainda as pessoas pegarem a demanda de análise fazer e devolver sem nem ter a crítica por quem está mais próximo do negócio, mas mesmo quem está mais próximo do negócio pode estar cheio de vieses. Talvez aquele algoritmo que eu dei o exemplo da Amazon talvez ninguém ia ter pensado nisso, olha, mesmo que tivesse pode ser que tenha rolado mil brainstorms dentro da Amazon antes de fazer isso e discussão com muitas pessoas inclusive com talvez mulheres no meio e ninguém pensando que poderia sair do outro lado um algoritmo que cuspsse só homens, porque você está pegando um histórico de dados da empresa e só tinha homens bem avaliados naquele mundo todo. Então eu entendo que um checklist sempre ajuda, pelo menos você vê os itens básicos. Estou olhando para isso, mas também o receio de checklist é isso aqui tem que ser um note, não não a caixa não a caixa, porque também isso vira uma um recurso para as pessoas falarem não, eu olhei o checklist cara, isso aqui é para te direcionar isso aqui é o básico do básico é para você conseguir pôr um pensamento crítico adicional, porque sempre vai ter muitas especificidades em cada cenário que você não vai conseguir abordar num checklist total então acho que tem que vir isso na caixinha com letras gritantes e olha, esse é o checklist básico do básico, mas seja crítico com o cenário específico que você está

Entrevistadora: ótimo muito bom Luciana, muito obrigada estou mandando aí no chat o form quando você puder, não preciso ser agora muito obrigada foi ótimo, mandou muito bem valeu

Luciana: eu vou sair porque estou atrasada para outra reunião.

Entrevistadora: Atrasada, né? Beijo, tchau

B.17

Mariana

Entrevistadora: Vai ler depois, tá? Uma coisa assim... ele tenta escrever em inglês o que você tá falando em português.

Mariana: É uma... mas a transcrição dele é bem boa. Fiz um teste outro dia. Teams está conseguindo acertar bem assim. Quando você, claro, seleciona o idioma certo.

Entrevistadora: Isso ele ainda não reconhece, né?

É. É, foi aqui.

Pronto. Mariana, assim, a ideia é que a gente consiga falar meia hora, mas fica à vontade se você quiser entrar em alguma questão, perguntar. Eu vou te contextualizar, né? Primeiro eu queria saber um pouco mais da sua experiência específica com algoritmo, tá? Se você pudesse me dizer, assim, de uma maneira geral, o que você já trabalhou, o que você tem mais, vamos dizer, domínio, né? E depois eu quero fazer, tipo, uma simulação com você.

Aí, como seria essa simulação? Eu vou pedir para você me dizer quais seriam as decisões que você tomaria para atingir, assim, dentro de um contexto que eu vou te contar, para atingir um resultado. E aí, se você puder, essa técnica que a gente usa chama Direct Storytelling, que é você, eu te dizer o que eu quero, né? Assim, passar uma situação hipotética e você, mas que é do seu dia a dia, né? E se você puder compartilhar sua tela e mostrar, assim, ah, eu primeiro pesquisaria o XYZ, usaria essa ferramenta, se você tiver que mostrar alguma ferramenta, quiser compartilhar a tela, fica à vontade também, tá? Mas aí vamos começar por isso, assim, queria saber um pouco mais da sua experiência. Sim.

Mariana: Então, Lari, eu me formei em estatística, né, na graduação, citei isso no questionário que eu respondi, e dentro da estatística a gente pode seguir para a linha de modelagem, né? De modelagem de algoritmos, que é bem próxima à ciência da computação. Então, foi a área que eu escolhi, inclusive me especializei em ciência da computação na pós-graduação, e aí dentro da faculdade a gente já tem contato com algoritmos, né, com programação, desde os períodos iniciais, né, que vem a parte de lógica, a parte de construção de condições para testar algoritmo, etc. Então, eu tenho contato com algoritmo já há bastante tempo, tá? As minhas experiências no mercado de trabalho são em duas frentes, uma delas é na parte financeira, né, eu trabalhei seguradora há muito tempo, então lá a gente tinha modelos de precificação de seguros, tinha

modelos de tomada de decisão para aceitação de seguro, também trabalhei com modelos antifraude na área financeira e na área de seguro também.

Então, lá eu tive uma experiência bem ampla nesse mundo, né, de banco e seguradora, até porque a seguradora que eu trabalhava também era banco. E depois disso, é, na consultoria o foco foi para indústria, né? Eu até citei isso lá no questionário que hoje eu já estou na indústria. E aí, de uma forma geral, em específico, né, nos últimos quatro anos, a atuação na área de segurança.

Então, é desenvolvimento de modelos, né, para a prevenção de acidentes, né, na área de segurança. Não só a prevenção de acidentes, mas também classificação de riscos, segmentação de riscos, então voltados para essa essa frente. Não sei se faz sentido, mas posso também citar alguns softwares que eu uso bastante.

Na faculdade, na experiência do banco, eu usava bastante o R, que é um software estatístico, né, que permite modelagem também. Atualmente, o que eu mais utilizo, né, é a linguagem Python. O que é próximo, tá? A gente consegue alcançar os mesmos objetivos utilizando uma ou outra linguagem de programação, mas hoje o Python é mais forte, até por conta da empresa que eu atuo.

Fora isso, a gente também trabalha bastante com linguagem SQL, mas aí já é outra parte, fora a parte algorítmica, né? Os algoritmos são desenvolvidos em Python.

Entrevistadora: Uma dúvida só que surgiu quando você falou de que você fazia modelo de precificação de seguros, né? Isso. Na sua outra experiência profissional, você, tipo assim, fazia um modelo para predizer se a pessoa era boa ou não, ou recomendada ou não, fazer um score em cima das pessoas possíveis para ter aquele seguro?

Mariana: Sim e não.

No modelo de precificação específico, a gente, com base nas informações pessoais, chegava no valor do seguro, diante informações como gênero, idade, localização de moradia, qual é o carro que a pessoa tem, o valor desse carro, né? Então, o preço do seguro era o resultado de um modelo que usava essas informações. E para a questão da aceitação, a gente também tinha um outro modelo, e esse modelo de aceitação era baseado em, essa pessoa tem histórico de acidentes com veículos, qual é o score financeiro dessa pessoa? Porque a gente, infelizmente, lida, na verdade lidava, não sei como está hoje, mas com muita fraude em seguro, né? De pessoas que contratam o seguro e em pouco tempo tem sinistro. Isso pode ser uma fraude.

Então, a gente utilizava informações que, na época, era permitido, né? Hoje, eu já não sei com o LGPD como é que está, mas tinha o score financeiro,

né? O score Serasa, a gente tinha acesso. E como a gente era banco, a gente também tinha acesso a algumas informações para quem era correntista, né? Ah, como é que é a experiência dessa pessoa, o relacionamento dessa pessoa com banco, é uma pessoa com relacionamento antigo, tem contratações com banco, como é que é? Então, isso era embasado para aceitação. Em alguns casos, que são muito poucos, a gente negava fazer o seguro, sim.

Entrevistadora: E essas regras, né? Assim, você poder, como você falou, na época não tinha LGPD, era outra história, mas essas regras de analisa o relacionamento, analisa a contratação, enfim, várias questões, você, quando começou a trabalhar com isso, você recebeu essas informações e você seguia ou tinham coisas que eles esperavam, né? O banco esperava que você sugerisse para melhorar a experiência.

Mariana: Então, lá a gente não trabalhava exclusivamente com sugestão do analista, a gente precisava provar com os dados. Então, assim, por exemplo, se eu sugerisse que entrasse na regra de aceitação a informação referente à localização de moradia da pessoa, eu precisava provar com os meus dados históricos que aquela localização tinha um número alto de sinistros.

Então, toda a decisão era baseada mesmo na informação e quanto a isso a gente respeitava. Claro que se fosse uma informação nova no mercado, por exemplo, algumas consultorias ofereciam dados como dados públicos, eu não vou me lembrar agora exatamente que tipo de dado, mas às vezes vinha uma consultoria oferecendo um dado público e a gente testava. Só que para testar eu não tinha histórico daquele dado público para dizer, poxa, realmente, dependendo da informação, tem mais sinistro ou não.

E aí a gente, às vezes, usava um grupo de teste, né, para avaliar aquele dado, mas assim, dificilmente ocorria uma situação onde o analista disse que essa informação tem que ser usada. Não, era mesmo comprovando com dados.

Entrevistadora: Tá, tá ótimo.

Tá legal, bem interessante, assim, né, porque, assim, tinha o direcionamento do analista, né, do negócio, mas antes de mais nada você tem que provar com a informação com dado mesmo.

Mariana: Sim.

Entrevistadora: Legal. Eu vou compartilhar aqui só para a gente ter uma, acho que às vezes lendo fica mais fácil, mas para você entender o nosso objetivo aqui, tá? Qual o meu objetivo? Eu gostaria de analisar, né, assim, de ter, de conseguir entender melhor quais seriam as suas decisões para fazer um modelo, fazer montar um algoritmo para fazer uma recomendação, é claro, é um algoritmo de recomendação especificamente, mas é para fazer uma divulgação de vagas. Tudo bem se você não tiver experiência com o algoritmo

de recomendação, né, especificamente, mas assim, você tem noção das tomadas de decisões que você faria e aí, assim, compartilhando mesmo. Qual o cenário? Vamos dizer que uma empresa de tecnologia de porte médio contrata uma empresa de divulgação de vagas, então, assim, uma empresa que está buscando um profissional, trabalha com tecnologia, trabalha com software, está buscando um profissional, que eu vou falar qual o perfil dele, contrata uma empresa para divulgação de vagas.

Então, a gente pode até citar um exemplo, dá para divulgar a vaga no Facebook, então você contrata o Facebook para fazer aquela divulgação de vagas, mas a nossa ideia aqui é assim, deixar mais pessoal, você trabalha numa empresa de divulgação de vagas, vem uma outra empresa de tecnologia que te contrata para anunciar a seguinte oportunidade, é uma vaga de tecnologia para profissional pleno, é para focado, né, que ele seja um desenvolvedor Full-stack ou uma desenvolvedora, com formação em engenharia de software, tecnologia da informação, de sistemas, coisas afins de tecnologia, uma experiência pelo menos de cinco anos e para trabalhar em regime remoto. Então, assim, essa empresa basicamente passa somente esses direcionadores, tá? Ela não falou mais nada, não falou de formação, especificamente.

Mariana: Onde a pessoa se formou, idade.

Entrevistadora: Isso, exatamente, não passou. E aí, passou isso, olha, eu quero, preciso de um profissional com esse perfil que eu tô te passando e que seja, né, assim, que tem uma conversão boa, né, com o fit, com o escopo do negócio, mas só isso. E falou, faz aí, tá? Então, a minha ideia é que você consiga falar, eu vou passar aqui para você ver, mas aí depois eu vou até parar de projetar para você e ir tirando da sua cabeça, tá? Fazer uma simulação, né, como você faria esse planejamento, a definição de requisitos, quais seriam os dados utilizados, escolher um método, né, qual método que você vai considerar para fazer esse algoritmo, como você faria.

Alguns dos desafios técnicos podem surgir, que você imagina, né, que aí você pode tentar contornar ou planejar. Como faria essa integração com uma vaga, né, com o sistema de divulgação de vagas da empresa. E os testes e validação, como você faria também, pode planejar um teste, como você fazia essas métricas de sucesso, né, assim, para dizer que funcionou ou não.

Feedbacks e interações. É basicamente isso. Se você preferir, eu vou indo ponto a ponto.

Mariana: Eu vou preferir sim, porque são muitos pontos e pode ser que eu me perca. Mas só para contextualizar, a ideia é desenvolver um algoritmo para encontrar um profissional para aquela vaga.

Entrevistadora: Isso, exatamente.

Que assim, tenha, você vai fazer essa divulgação, o anúncio vai correr em alguma plataforma, aí a gente não está fechando, né, assim, você pode definir aí um cenário hipotético, mas vai correr, esse anúncio vai chegar nas pessoas, de alguma forma, que você acha, né, que o algoritmo vai buscar, que sejam realmente o perfil do que você está buscando, que tenha a fit.

Mariana: Tá bom. Achei difícil, mas vamos lá.

Definição de requisitos. Como você definiria os requisitos para o algoritmo? Bom, vou pensar numa situação aqui hipotética, eu não conheço dados onde eu buscaria dados como esses, mas eu vou pensar que pudesse ser o LinkedIn, tá? Até porque eu sei que lá no LinkedIn é possível você fazer anúncios de vagas e, muito com certeza, eles têm um algoritmo para isso. Então, suponhamos que a minha base de pessoas, de profissionais, esteja no LinkedIn, né? Eu preciso conhecer essas pessoas.

E aí, definição de requisitos. Como você definiria os requisitos para o algoritmo? Bom, eu diria que eu preciso de pessoas com a formação na área que eu especifiquei, né? Tecnologia da informação, desenvolvimento de softwares. Eu, se não tivesse essa informação explícita, eu pegaria o ano de formação dessa pessoa para saber se, de fato, ela tem um período de formação superior a cinco anos, o que não necessariamente se refere à experiência, mas se eu não tiver lá anos de experiência, eu usaria o ano de formação dela para entender quanto tempo que ela tá atuando, né? Eu veria também a questão de tempo e empresa, porque pode ser que a pessoa tenha se formado, mas pode ser que ela não esteja trabalhando, né? Na área.

Então, essa informação poderia ficar incorreta na hora de calcular a experiência. Então, no LinkedIn eu sei que é possível você saber onde a pessoa trabalha, né? Então, veria isso também. Você mencionou experiência, formação, tá? Ah, e regime remoto, poderia ver se ela já trabalhou em regime remoto ou não, mas como o regime remoto normalmente é o mais requisitado, acho que quanto a isso eu tô tranquila.

O que aparecer, acho que a pessoa vai ficar confortável escolher regime remoto. Então, a gente, nada, pode voltar para a definição e requisitos. Como você define os requisitos para esse algoritmo? O que eu falei, na verdade, seria o método de busca, né? Na verdade, o que eu falei foram os dados utilizados.

Eu usaria os dados de LinkedIn, baseado nessas informações que eu mencionei. A definição de requisitos, eu acho que eu tô com um pouco de dúvida no que seria exatamente.

Entrevistadora: É, do tipo, como você começa a fazer, a gente pode até falar de regras, assim, né? Para o algoritmo, para montar ele, assim.

Então, você falou dos dados, aí, a partir desses dados.

Mariana: É, eu falei dos dados agora.

Entrevistadora: É, você também falou alguns requisitos, né? Do tipo, o tempo que ele se formou, o tempo, que são os dados, mas que estão ligados àquelas regras que a gente vai definir.

Tá, tá bom. Aí, você me diz, assim, por exemplo, você quer, quando você falou, vou olhar o tempo de experiência, o tempo de permanência numa empresa. Aí, você já foi falando, mais ou menos, por quê? Para saber se a pessoa ficou, né? Assim, aí, isso também é um requisito, assim, que você pode colocar.

Mariana: Ah, tá. Eu posso entender, também, a definição de requisito como a construção de variáveis, né? De informações com base naquelas que eu tenho, né? Não somente, ah, o ano de formação, mas do ano de formação, o que a gente tira? A experiência da pessoa. Tá bom, entendi agora, claro.

Então, os requisitos que eu poderia adotar, né? Seria essa construção de informações além, como eu já mencionei. E, pensando em regra, também, de algoritmo, eu poderia especificar que eu quero ver pessoas, né? Profissionais que atendam, sei lá, 70

Então, eu posso adotar como requisito, ah, um match superior a 70

Entrevistadora: E, aí, até pensando, assim, a empresa te passou aqueles dados, né? Assim, essa premissa, né? Dessas questões aqui, dessas características, mas, assim, você, né? Pensando nesse cenário que você trabalha em uma empresa de divulgação, né? E vai desenvolver esse algoritmo, você quer ter o melhor resultado possível. Às vezes, assim, você fala, bom, preciso de mais informações ou preciso focar em alguma informação específica que eu tenho certeza, vou por esse nicho aqui, que eu tenho certeza que vai ter mais resultados.

Então, você tem as suas inferências também, pode ter ou não, e, aí, você pode colocar também para fazer um filtro, enfim, para você achar que vai melhorar a experiência desse algoritmo, tá?

Mariana: Ah, tá. Um ponto que é legal, mas, assim, se eu tivesse desenvolvendo esse algoritmo, eu faria depois de rodar ele, né? De desenvolver ele a primeira vez, seria para entender o perfil das pessoas que chegou. Então, das pessoas lá que têm, sei lá, 70

Entrevistadora: Sim, tá legal.

Mariana: Respondi essa...

Entrevistadora: Respondeu, respondeu. E aí, assim, já pensando no desenvolvimento do algoritmo em si, né? As questões técnicas mesmo, se você tem algum método que você acha que funcionaria para colocar, assim, de ár-

vore, enfim, qualquer um que você tenha preferência, achando que vai dar esse direcionamento legal.

E, assim, possivelmente, a gente tá falando de simulação, se surgiriam alguns desafios técnicos que você já imagina por experiência que pode surgir, e como você faria essa integração no sistema mesmo, assim, né? Da divulgação.

Mariana: Tá. Um método que eu escolheria, né? Pensando nesse critério que eu mencionei, que seria de dar um score para a pessoa, que seria o match dela com o que eu procuro.

A árvore é um grande método, tá? Ele é muito bom para esse tipo de classificação, porque a ideia é classificar o profissional, né? Do quão ele me atende, tá? Pensando naquela descrição de vaga, né? Então, a árvore de decisão, ela ajuda, porque cada raminho da árvore seria uma variável, né? Uma feature desse modelo, e aí suponhamos que a feature seja ano de formação, aí eu quero saber se atende ou não. O outro lá era graduação que a pessoa fez, será que atende diante das que eu escolhi? Então, a árvore seria um bom método para classificar esse profissional entre, ah, ele tá apto ao que você escolheu ou não.

Desafios técnicos. Quais seriam os principais desafios técnicos enfrentados? Olha, um desafio técnico seria base de dados, né? Porque para que eu possa criar um algoritmo como esse, o ideal é que eu tenha uma base de dados que mencione diversas pessoas e tenha a resposta ali, se ela é apta ou não. Mas eu poderia construir essa base. O negócio é que muitas vezes não existe tempo hábil de você construir uma base de dados para treinar o modelo.

Então, um desafio técnico seria, tá bom, Mariana, você quer desenvolver um algoritmo para classificar novos profissionais entre aptos e não aptos ao seu job offer. Como você vai treinar esse modelo? Como você vai ensinar ele o que é ser apto ou não? Isso é um desafio. Por quê? Porque eu posso construir essa base, mas aí eu vou construir essa base com base no que sai da minha cabeça.

Tipo, chegou uma pessoa lá na base. João, formado em tecnologia da informação, se formou em 2015. Aí eu vou dar resposta, né? Eu não sei se está claro para você.

Entrevistadora: Tá, tá.

Mariana: Aí que bom, porque eu preciso construir essa base para treinar o meu modelo e automatizar isso. E aí eu acho que um dos maiores desafios seria ter essa base de treinamento.

Entrevistadora: Sim, e até, se você quiser, até você estava citando um exemplo, João e tal. Se quiser ir por dentro disso, eu acho que é até interessante também, assim, te perguntar já. Como você falou, isso sairia da minha cabeça.

Se fosse uma situação real, assim, até tentando pensar na sua realidade hoje, o que que você faria para ter um apoio nesse desenvolvimento? Então,

assim, para tentar diminuir essa inferência, né, que você faria?

Mariana: Olha, sempre que eu passo por uma situação como essa, onde uma decisão precisa ser tomada para continuar uma análise, eu contato um colega par. Normalmente um par, né, que tem ali o mesmo cargo, a mesma experiência que eu, para que a gente possa, enfim, discutir, tomar uma decisão que seja concordada por mim e pela pessoa. E aí, assim, eu não sei se você quer chegar mais a fundo, por exemplo, lá tem relação a gênero, tem relação a... Não, tá? Em todos os lugares que eu passei, assim, o que eu sempre me preocupei foi em discutir com pessoas que fossem da mesma área que eu, independentemente de qualquer coisa, entendeu? Tem lugares onde a gente, dependendo da empresa, né, que a gente vê que muitos profissionais, às vezes, não são do mesmo gênero que você, né, mas eu nunca enfrentei nada que pudesse me fazer desconfortável, tá?

Entrevistadora: E aí você... Mas você busca pessoas que tenham a formação técnica, como a sua, para você trocar.

Você acha ou já fez, né, que chamar uma pessoa do negócio em si, que seja especialista no negócio, isso você já fez? Isso ajuda?

Mariana: Então, com relação ao negócio, eu sempre envolvo, mas não nessa parte técnica. Por quê? Porque às vezes a pessoa pode não entender o que está acontecendo. Mas aí eu volto até o exemplo que eu dei, que é, quando eu rodo a primeira versão desse modelo, que normalmente não é utilizada, tá? É só a primeira versão, o primeiro resultado.

Eu chamo alguém que entenda do negócio para dizer, faz sentido? Faz sentido que a gente chegou, por exemplo, no mundo dos acidentes? Faz sentido o modelo dizer que no turno da manhã tem mais acidentes do que no turno da tarde? Porque às vezes essa informação, para mim, é super relevante no modelo, mas o cara vai dizer, não, faz sentido, a equipe é a mesma. Então, eu envolvo bem, sim, a pessoa do negócio, mas assim, pós ter ali um primeiro resultado.

Entrevistadora: Entendi, tá. Até para saber se está alucinando ou não.

Mariana: É, está alucinando ou não.

Entrevistadora: Tá, maravilha. E você imagina como faria essa integração?

Mariana: A integração? Como o algoritmo foi integrado no sistema de divulgação?

Entrevistadora: Isso depende mais do... porque a gente não deixou fechado, né? Assim, você falou que você faria no LinkedIn, mas aí é uma questão técnica que realmente não faz muita... não é muito relevante, né? Assim, para isso.

Mariana: É, e aí, assim, como eu até te falei, eu... É o mundo... seria um mundo novo para mim. Então, como eu implementaria? Eu acho que, pensando no LinkedIn novamente, tá? Eu acredito que lá no LinkedIn, quando você é pro, premium, sei lá, você pode... e é do ramo de recrutamento, você pode preencher algum formulário, alguma coisa que possa indicar as características que você procura.

Mas isso já seria o modelo deles, né? Agora, integrar o meu modelo... olha, no mundo ideal para mim, assim, seria bom que esse site de divulgação, esse sistema de divulgação, tenha um método de API, né? Para que eu possa consultar esse dado. Aí eu teria lá o dado de todo mundo. Então, depois de eu rodar meu modelo, eu vou chegar nas pessoas que eu quero.

E aí, chegando nas pessoas que eu quero, eu faria uma consulta nesse sistema, ó. Me passo o currículo dessas pessoas.

Entrevistadora: Tá, entendi.

Mariana: Talvez isso. Se bem que você quer criar um anúncio, né?

Entrevistadora: É.

Mariana: Criar um anúncio. Difícil. Porque, assim, a gente já tem informação lá para requisito. O anúncio poderia usar aquilo.

Entrevistadora: Entendi, tá. Aí já, assim, o anúncio, na verdade, seria assim, o primeiro contato, né? Porque aí você já selecionou a pessoa e você vai lá e pega a pessoa.

Mariana: Já teria um filtro. E as pessoas que não interagiram com o meu anúncio, pode ser que elas sejam também profissionais que atendam ao que eu quero, né? Tá. Mas aí que eu tô falando, seria um mundo ideal, porque a gente não sabe se esses sistemas de vagas, eles disponibilizam os dados de todo mundo, né? Não sei nem se isso é permitido.

Entrevistadora: É, hoje em dia é mais complexo, né?

Mariana: Ah, então eu acho que é isso mesmo. A gente bota o anúncio, aí eu botaria aquilo lá, que tá lá. E veria quem fosse interagir com aquele anúncio.

Seja curtindo ou seja se inscrevendo no link. Aí eu trabalharia com essas pessoas.

Entrevistadora: Entendi.

Mariana: Porque eu acho que seria permitido, né? A pessoa, geralmente, quando ela segue o link de uma vaga, ela meio que autoriza que a pessoa veja seus dados.

Entrevistadora: Mas aí quando você faz, nessa situação de fazer o anúncio, eu imagino, né, que você vai falar assim, mostra para as pessoas do perfil tal, tal, tal. É, sim.

Mariana: Mostra para as pessoas do perfil tal, com a formação nas áreas X, Y e Z, com tempo de formação superior a cinco anos. Essas coisas são fáceis de mencionar, sim.

Entrevistadora: E você colocaria alguma coisa além daquilo? Porque você falou ali antes, né, que a gente pensou assim, tipo, primeiro, né, em buscar as pessoas direto e agora a gente tá falando do anúncio.

É. Você chegou uma hora e falou assim, depois que retornassem essas pessoas eu ia olhar para ver se não tá numa bolha, porque isso não é bom para a empresa. É. O que seria essa bolha? E se você, pensando agora nessa situação de anunciar, de buscar, fazer um, pescar essas pessoas, você teria alguma, um direcionamento para isso? Para não retornar a bolha, né?

Mariana: É, sim. O que eu acho importante como recrutador, né, na área de tecnologia, é que você diversifica um pouco as pessoas que você vai entrevistar.

Por quê? Não necessariamente um profissional que vem de uma universidade particular, ele não é competente para atender aquela vaga, né? É claro que quando a gente vem de faculdades federais ou faculdades com nome, a gente tende a entender que profissionais desse mesmo, dessa mesma bolha, vão atender o que você procura. Mas eu, Mariana, eu não fecharia os olhos para entrevistar pessoas de outras faculdades, de, né, outras formações. Eu prezo muitas essas coisas da formação porque eu não, eu não costumo, eu não costumo pensar que, por exemplo, diversificar em relação a gênero pode não, eu posso não, eu posso acabar trazendo uma questão de, ah não, vou preferir entrevistar mulheres com a formação.

Não sei, tá? Eu nunca passei por essa coisa de recrutar, mas eu já imagino que, naturalmente, o número maior de pessoas que atendam a minha vaga seja de homens, né? Possivelmente vai ter muita gente também de formação, eu acredito, tá? De formação federal, porque a área de tecnologia, né, ela é bem forte nas universidades federais. Agora tá aumentando nas particulares, mas pelo que eu sei, por exemplo, estatística até até 2021 não tinha curso de graduação em estatística e em faculdade particular, só nas federais. Enfim, acabei viajando aqui.

Entrevistadora: Não, pode ir.

Mariana: Mas eu acho que eu prezaria mais no quesito formação e, claro, na hora de selecionar os currículos, vamos supor que venham 200, não dá para entrevistar 200 outras pessoas. Vamos fazer um mix aqui entre gênero, idade, entrevistar pessoas com idade mais jovem, mas também com idades, né, mais avançadas, porque tem pessoas que fazem migração de carreira e tá tudo bem.

Eu acho que eu prezaria muito mais para a questão da formação, batendo

a tecla na formação. Ah, beleza, a pessoa formar em TI, mas ela trabalha em TI? Ela tem cursos fora, né, nessas plataformas gratuitas ou não. Enfim, eu bateria muito nessa tecla assim, tá, de formação.

Entrevistadora: Tá. Então, até continuando aqui, na parte de teste, né, você já falou um pouco, mas assim, até para entender, você falou da métrica de sucesso, né, se eu colocarei um score, né, e ter um match, mas se quiser falar um pouco mais disso. E aí, a partir daí, você falou assim, ah, primeiro tem que fazer uma base e aí eu vou começar a discutir em cima disso, vou chamar as pessoas de negócio e tal.

Como seria esse feedback que você pode pegar e fazer a interação, né, em cima do modelo para fazer alguma melhoria?

Mariana: Aham. Bom, o que eu faria, né, depois da primeira versão e depois de discutir com pessoas que sejam do negócio também e também que seja, por exemplo, a pessoa que quer esse profissional, que vai trabalhar junto com esse profissional. O perfil que eu encontrei foi esse, você acha que faz sentido entrevistar esse perfil, dar uma diversificada em relação à idade, em relação ao gênero, porque não necessariamente, né, como eu falei, pessoas com mais idade têm mais experiência que pessoas com menos idade, isso aí.

Então, com relação ao teste, não sei se eu faria o teste AB, mas eu faria uma análise exploratória em relação ao resultado desse modelo, que foi o que eu te falei, para entender o perfil das pessoas que o modelo indicou como altas chances de dar match com o que eu quero. Quais métricas você usaria para avaliar o desempenho? Então, a gente tem algumas métricas para avaliar desempenho de modelo que seriam em relação à divisão da base de dados em treino e teste, né. Então, eu uso parte dessa base para treinar o modelo e dizer, olha, é sim, né, para a pessoa que dá o match, que tem essas informações, aí na base de teste eu só tacho o modelo para ele scorar e avalio no olho se ele acertou ou não, porque, como eu te falei, eu que estaria dando essa flag, né, eu que construiria essa base de treinamento.

Mas aí, de novo, olharia a distribuição dessa base teste classificada pelo modelo, entenderia se de fato, olha, ele chegou num legal. Todas as pessoas que estão aqui, ele indicou como deu match e, de fato, elas dão, porque elas, olhando assim, né, elas têm a formação que eu quero, tá bom? Feedback, como avaliar o feedback interno e externo que podem levar a modificações? Feedback. Olha, eu acho que seria interessante, fugindo um pouco até dessa parte mais técnica, para a empresa que eu contratei, que foi a empresa que é de recrutamento, é entender como o profissional de recrutamento se faz sentido aquele resultado.

Porque, às vezes, o modelo chega num lugar que não necessariamente a gente, como desenvolvedor, entende. Então, por exemplo, vamos supor que o

modelo também indicou que deu match uma pessoa que não tem a formação na área de tecnologia, mas sempre trabalhou em empresas de tecnologia, faz cursos na área, mas, sei lá, suponhamos que essa pessoa seja de humanas, fez faculdade, nada a ver, mas estuda sobre ciência de dados ou sobre engenharia de dados. De acordo com a experiência do recrutador, olha, você já entrevistou pessoas nesse perfil? Deu match, foi bom?

Só para eu saber se o modelo está me direcionando pessoas, de fato, que atendem a... Eu não sei se eu tô viajando muito, tá?

Entrevistadora: Não, tá nada viajando, você tá ótima.

Mariana: Eu tô preocupada, assim, de caramba, tô falando um monte de abobrinha.

Entrevistadora: Bom, eu tô gostando e tô aprendendo bastante, não tô achando nada viajante, tá ótimo.

E você tá construindo, tá fazendo um exercício, construiu o pensamento, é exatamente isso que eu quero, assim, que é importante. Tá, então, eu falei. Não, perdão, pode falar.

Mariana: Não, só para concluir, eu acho que o feedback é bom ter de pessoas não técnicas, porque o feedback de alguém não técnico que entenda do negócio vai me direcionar se o modelo está tendo uma assertividade legal, né, uma curação legal ou não.

Entrevistadora: Sim, ótimo. Aí, os resultados e impactos, né, assim, o impacto na contratação, primeiro, efetividade, né, como você pode avaliar a efetividade do algoritmo na recomendação de candidatos com mais ou menos que os desejados.

Você já falou bastante disso, mas essa parte do impacto na contratação, de fato, quantas pessoas você selecionou e quantas pessoas foram contratadas, né, assim, isso seria algum direcionador para você ou isso retorna, né?

Mariana: Com certeza, se eu, trabalhando nessa empresa de recrutamento, desenvolvo um algoritmo e eu, por alguma razão, descubro depois que dois, sei lá, 20 profissionais que eu indiquei para as empresas, mais de 50

Entrevistadora: Sim, com certeza. Não, mas assim, a minha ideia é que é pensar isso mesmo, assim, foi ótimo, porque você foi construindo pensamento e às vezes a gente vai e volta e pensa e aí é ser uma coisa bem espontânea mesmo, sabe? E, assim, seria o que você realmente pensaria e como planejar, como fazer, como executar. Eu acho que é bem isso mesmo, Mariana.

Deixa eu ver se tem mais, acho que acabou aqui. Acabou. Mas, assim, tem alguma coisa que você acha que você não falou, que eu não perguntei, que traria alguma ou outro direcionamento?

Mariana: Olha, eu te falei, né, que eu nunca fiz algoritmo de recomendação, então isso é novo para mim, mas eu tenho certeza que depois que eu fizesse a primeira versão, muitas coisas mudariam, porque, de fato, esse desafio da base de dados de treinamento, para mim, é o maior, tá? O que é uma pessoa apta ou não? Que exclusivamente atende àqueles requisitos e se as pessoas que atendem àqueles requisitos, da forma que eu especificar, de fato, são pessoas que estão aqui nessa bolhinha, né? E a gente não tá olhando o resto, o todo.

Então, isso seria um desafio legal, assim, de estudar. E até que ponto vale usar informações pessoais também, né, para treinar esse modelo? Eu não mencionei, eu não usaria informações como gênero, por exemplo, mas se eu usar, pode ser que o modelo me retorne um resultado mais diversificado, porque ele vai conhecer perfil de homens e de mulheres, né? Se eu não cito gênero como uma feature, né, como uma variável de treinamento, ele simplesmente vai ignorar esse fato. Então, ele... porque, assim, lembrei de uma coisa legal.

Quando a gente treina o modelo, a gente pode usar esse tipo de informação e balancear meu dado. Suponhamos que, na base que eu uso para treinar, tem 80

Eu quero que ele balanceie isso, eu posso dizer, olha só, eu quero que a coluna, né, que a variável gênero fique balanceada. Então, ele vai dar um jeito, né, de conseguir balancear, e na hora de me retornar, ele me retornar 50 a 50, sabe? Mesmo que no aprendizado dele tinha 80 a 20.

Entrevistadora: Sim.

Mariana: Esse tipo de coisa é legal pensando em incorporar diversidade no retorno desse algoritmo.

Entrevistadora: Mas você, com esse input que você teve da simulação, você faria isso desde o início ou você esperaria retornar um resultado para depois pedir para alguém do negócio te ajudar, ou você mesma já colocaria como um direcionamento?

Mariana: Não sei. Eu acho que depende.

Eu faria por conta própria caso fosse um desbalanceamento muito forte, tipo 80 a 20. Se eu vejo que 80

Mas eu acho que se fosse uma diferença muito grotesca, vale sim o balanceamento, até porque isso, tirando o tema em questão da frente, isso é recomendado para você quando você vai treinar algoritmos. Não é interessante que uma informação seja muito desbalanceada, porque o algoritmo não vai conseguir aprender legal. Se eu tenho só 20

Então, balancear é tecnicamente recomendado.

Entrevistadora: Entendi. É isso, Mariana, muito bom.

Eu gostei do jeito que você conduziu aí. É exatamente isso, assim, quanto mais espontâneo melhor e você trouxe muitas informações técnicas boas, acho que vai ser ótimo.

Mariana: Que bom, que bom, Lari.

Eu não sei se ficou confuso, mas caso você, enfim, queira se aprofundar melhor e tal, eu fico de disposição. Foi legal poder pensar em coisas que eu nunca tinha pensado antes. De fato, isso nunca entrou na minha linha de trabalho, assim.

Nem de pesquisa, na época da faculdade, eu nunca trabalhei com recomendação mesmo, mas foi legal pensar nessas possibilidades.

Entrevistadora: Muito bom, muito obrigada, viu?

Mariana: De nada. Qualquer coisa, pode chamar. Valeu, Lari.

Entrevistadora: Tchau, tchau.

B.18

Rebeca

Entrevistadora: Então, entre a pessoa e o dado específico. Tá, deixa eu ver se tá com a transcrição certa, porque às vezes vai errada a linguagem, aí é um probleminha. Cadê? Fala alguma coisa aí, Rebeca, deixa eu ver, vamos testar, testando.

Rebeca: Vamos ver se tá gravando.

Entrevistadora: Acho que tá gravando aí, né? Apareceu?

Rebeca: Aqui tá. Tá.

Entrevistadora: Então, vamos começar. Eu compartilhei a tela aqui, a minha ideia é o seguinte, eu quero... Você falou, se eu falar besteira. Se você falar besteira, eu não vou saber, porque eu não sei tecnicamente.

Então, fica de boa. A questão que eu quero analisar aqui é mais assim, os caminhos do pensamento, assim, o que que você vai decidir, vai fazendo, vai montando para decidir para uma solução específica que eu vou te passar. Mas a ideia é você, tudo que você pensar, assim, em relação à solução, vai falando, que aí a gente pode falar e depois falar assim, acho que falei errado, vou mudar, não tem problema, tá? É isso mesmo que eu pretendo observar, tá? Aí, no começo, se você quiser, também, peraí que... Eita, sumiu tudo.

Se você quiser... Tá vendo minha tela ainda? Tá preta. Agora foi para a tela da Deloitte. Deu um pico aqui que apagou a tela.

Se você quiser compartilhar alguma coisa, sei lá, alguma ferramenta que você queira mostrar, alguma coisa assim, fica à vontade. Mas eu tô compartilhando só para ir ajudando, para saber onde que tá, né, assim, pautando a nossa conversa. Aí, primeiro, eu queria uma contextualização, se

you can give, about you, more the academic or professional that you have had contact with the algorithm, have had that to develop and what is your position today, in your current work, né? Position that I say, well, that you are developing.

Rebeca: Tá, só ela falando assim, desde a gente tá dando um eco.

Entrevistadora: Deixa eu fechar, então, peraí.

Rebeca: Aí, eu fico escutando, mas eu vou ficar igual a mulher de sanduíche, ixi.

Entrevistadora: Tá ouvindo ainda? Mas eu fechei aqui. Deixa eu ver.

Rebeca: Não, acho que agora parou o eco. Você fala assim, para eu falar a minha formação e como que eu comecei a trabalhar, fazendo um resumo, coisas assim?

Entrevistadora: É, mas, assim, voltado mais para quando, nas experiências que você teve, que você desenvolveu o algoritmo, você teve, né, não precisa dizer o que é específico, mas, assim, fazer um passado geral.

Entrevistadora: Tá, eu sou formada na Unesp, em estatística, lá a gente pouco tinha acesso ao computador, eu não tinha computador na época também, até por questão financeira e tudo mais. Lá a gente teve uma matéria, uma disciplina que era de Pascal.

Foi a primeira vez que eu comecei a ter contato com qualquer coisa de programação. Depois disso, eu fui para o Itaú, comecei a trabalhar no Itaú, e lá eu comecei a utilizar o R, fazia previsão, então usava o R no dia-a-dia, no comum. Depois do Itaú, eu fui para lojas americanas, que eu comecei a ter um pouquinho mais de contato com o SAS, desenvolvendo modelos de predição também, e também de vendas, né, que eles trabalhavam muito com, eles chamam de eventos, que era Páscoa, Natal, enfim, tudo mais desse tipo.

Antes de ir para americanas, eu fui para trabalhar na Oi, que eu também fazia modelo de previsão, era usar uma regressão linear. E aí, depois eu trabalhei na Losango, na Losango foi quando eu comecei mais ainda a ter contato com a programação, que eu não fazia mais modelo estatístico, mas eu usava algoritmo para tratar base de dados e trazer análise para diretoria e tudo mais, mas não necessariamente modelo estatístico, era mais usando programação mesmo. Então, foi aí que eu comecei mais a programar.

E aí, depois, foi que eu caí aqui na Vale, e aí que eu tive que me jogar, principalmente no Paito, porque então eu usava o SAS, o R, na faculdade, apesar de a gente ter tido o Pascal, a gente usava muito o R e o SAS, e aí aqui eu comecei a ter que me jogar já direto no Python e desenvolvendo modelo e chora, tipo isso, senta e chora e vai. Não dava tempo de pensar lá, ele tinha que estudar. Eu fui quando eu entrei, estava bem na época do Covid, como que

eu vim parar na Vale? Foi meu professor do MBA que me trouxe, e aí eu não podia falar, tô com dúvida, eu tinha que simplesmente, eu tinha que entregar o coronavírus.

Ele me chamou junto com a Liriele, que era uma outra cientista de dados muito boa, e mais um menino também da IBM, e aí a gente ficou trabalhando, desenvolvendo, tentando fazer previsão para o Covid, que eu acho que foi um grande erro, porque, cara, sinceramente era uma coisa impossível de se prever, mas a Vale foi reconhecida por ter desenvolvido, ele ganhou vários prêmios por conta disso, hoje até ele tá lá na TIM, porque ele alavancou a carreira dele em cima disso. Mas eu, particularmente, olhando os resultados, assim, como estatística, eu falo, era quase que jogar a moeda para cima e falar verdadeiro ou falso, e aí tipo, acerto de 50

E aí, depois, eu fui para um projeto do RH também, fazendo análise, seguindo o modelo, puxando mais para estatística, mais tranquilo, e depois agora aqui também, que tá mais a minha vibe, não é aquela coisa automa- tizada, aquela coisa que tem que jogar em produção, né, tem que estar tudo automatizado, tudo, o mais, o código tem que estar o mais enchuto possível, pensando na velocidade que ele vai demorar para executar e tudo mais, então, agora, tô usando algoritmo, mas de uma forma mais, mais na vibe analítica. Era isso que era a ideia?

Entrevistadora: Isso mesmo, é isso aí. E aí, eu vou te passar um cenário fictício, para você me dizer qual solução você daria, a gente vai por partes, tá, eu vou guiando ali.

Imagina que uma empresa de tecnologia de porte médio, contrata uma empresa de divulgação de vagas, e você trabalha nessa empresa de divulgação de vagas, tá, então, uma empresa de tecnologia, tá querendo, tá com uma vaga aberta e tá querendo divulgar a vaga aberta que tá, e aí ela te contrata, né, você tá lá na vaga de, na empresa de divulgação de vagas, para construir um anúncio da seguinte oportunidade, é uma vaga de tecnologia para profissionais pleno, com skills para desenvolvedora ou desenvolvedora full stack, com formação em engenharia de software, tecnologia da informação, sistemas e afins, uma experiência de cinco anos e regime remoto. Só para explicar melhor, esse anúncio, imagina que assim, você tá no Facebook, no Instagram, e aparece um anúncio, aquele anúncio, com certeza, tá aparecendo, porque já leram o seu perfil, entenderam que você vai dar match com aquele anúncio, então, a ideia é você entender, aí, assim, a plataforma específica não precisa ser até isso, sabe, se for no LinkedIn, no Facebook e tal, tanto faz, mas assim, quais dados profissionais da pessoa você gostaria de ter, para modelar esse algoritmo, para fazer com que ele apareça, o anúncio, né, a publicação apareça, para aquelas

pessoas que provavelmente vão querer se candidatar.

Rebeca: Entendi.

O que que eu quero puxar de variável, para que eu vou analisar e vou falar, cara, isso aqui é interessante para poder, então, se a gente quer pegar uma pessoa que é...

Entrevistadora: Só passar para a próxima fase, que aí você, aí depois eu volto nesse slide, que aí você responde por partes, que aí fica mais fácil. Aí, por exemplo, qual as... Na parte, na fase de planejamento e concepção, como seria a definição dos requisitos, a partir daqueles requisitos que eles passaram, né, e quais seriam os dados utilizados?

Rebeca: Tá, perguntando aqui, né, essas seriam as variáveis que eu vou analisar para poder, tá, pensando... Primeira coisa, né, a gente quer alguém que tenha conhecimento em... Vaga de tecnologia, a gente quer alguém que tenha conhecimento em programação. Puxaria a formação acadêmica, tanto, né, se é tanto graduação como pós-doutorado, como fala? Pós-graduação, né, tanto MBA como mestrado, doutorado e tudo, mas eu também puxaria o histórico de... de experiência, porque tem muitas pessoas que, às vezes, estão formadas, né, por exemplo, tava lá no projeto que é de geotecnia e o menino era biólogo e ele estava trabalhando como cientista de dados e ele tinha um cientista de dados foda.

Então, eu acho que só a formação, ela não leva a trazer, às vezes, a melhor pessoa, né, pode ser que a pessoa ela tenha tido uma experiência muito maior do que a formação em si. Então, o histórico, o histórico... histórico da carreira mesmo, então, profissional, né, histórico profissional dessa pessoa. É, deixa eu pensar, talvez programas que ela usa, né, que ela tem conhecimento, conhecimento em... tem conhecimento em R, em SAS, XPTO.

A experiência está pedindo cinco anos, então, logo quando eu pego esse histórico do profissional, eu consigo trazer o tempo que ele tem de experiência também. É, por ser regime remoto, não faz diferença o lugar, né, o lugar que é minha empresa e a localização que essa pessoa vai estar. Acho importante, talvez, trazer conhecimento de... de inglês, porque muita coisa quando a gente está usando é inglês, então, uma pessoa que tem conhecimento inglês facilita muito, tanto para gravar vídeo, para conseguir, às vezes, a pessoa não tem necessariamente... ela sabe fazer alguma coisa, mas ela vai buscar livro, vai buscar livro, né, na internet, vai buscar vídeo, vai buscar as coisas, tem muitos que são em inglês.

E aí tem aquela outra parte, né, eu acho que não só aquela pessoa que ela já tem o conhecimento está preparada para tudo, mas é que aqui está pedindo por experiência, mas... porque, às vezes, a pessoa ela não tem... não tem o

conhecimento, mas ela tem tudo, os skills para se preparar para ter, né, é aquela pessoa que tem a... só que aí não faria sentido numa vaga de entrevista, né, para buscar essa variável. Estou pensando aqui, lá, falando besteira. Eu não tenho como ver se é um perfil ousado, não sei se esse tipo de coisa eu consigo ter no... criatividade, ousadia, essas coisas dentro do perfil de um Instagram, Instagram, não, de um Facebook, ou então de um LinkedIn.

Tá, são coisas mais abstratas, né? Exatamente, é, eu acho que essas coisas... para trazer um... tipo, eu estou pensando no algoritmo, tentando pensar no LinkedIn, assim, ele vai tentar associar coisas que você já trabalhou ou coisas desse tipo, né, coisas que você está buscando, coisas que você tem informação, alguma coisa mais relacionada. Eu acho que fica mais abstrato as outras coisas, que eu acho que vai muito mais além.

Entrevistadora: Tá, sim, tem razão.

E aí, a parte do desenvolvimento em si, você pensou ali nos dados, nos requisitos, né, no que você quer puxar das pessoas, qual método você escolheria, ou alguma técnica de machine learning, não sei, tanto faz, assim, o que você considera para fazer esse algoritmo, e em relação a desafios técnicos, você considera que pode ter algum desafio técnico, que isso até ajude a nortear a escolha do seu método, né? E aí, assim, basicamente isso, pode ficar nessas duas.

Rebeca: O que eu escolheria, sem dúvida, seria trabalhar hoje, hoje, né? Antigamente eu diria que seria o SAS ou o R, hoje eu diria que qualquer coisa eu tentaria puxar para o Python, pela disponibilidade que ele tem.

Entrevistadora: Mas qual método que você faria no Python? É, tipo, você escrever o código, mas tem alguma coisa que você ia seguir, alguma coisa específica, assim, tipo, não sei, algum algoritmo já modelado?

Rebeca: Eu ia pegar algum modelo, algum modelo, é, um modelo estatístico, de, não sei dizer uma, assim, teria que pensar com mais calma, Lari, mas eu partiria primeiro pensando no Python e, assim, ah, que modelo que eu vou usar, que variável, depois que eu tenho essas variáveis que eu iria utilizar, ah, como que eu posso usar da melhor forma essas variáveis? E aí pensar em algum modelo, provavelmente eu ia usar algum modelo de regressão, algum modelo predictivo, alguma coisa nesse sentido, mas eu não sei te dizer, assim, rápido, precisaria pensar com mais calma.

Entrevistadora: E tem algum desafio que você acredita entre a captura dos dados, a análise dos dados e gerar esse modelo para fazer essa classificação, né, o padrão, enfim?

Rebeca: Para mim, totalmente, porque como eu não tenho conhecimento da engenharia, para mim seria totalmente um desafio, mas, assim, coisas que

seria super possível ter uma pessoa que tem o conhecimento, por exemplo, Tadeu, eu acho, Fernando, são pessoas que, por exemplo, que eu acho que conseguem fazer esse link, que não é o meu perfil. Então, para mim, eu não sei se é nesse sentido que você está perguntando, ah, desafio, porque tanto tem o desafio técnico mesmo, que, ah, então, a pessoa tem o conhecimento e, mesmo assim, é difícil executar, como tem o meu caso, que é, eu não tenho conhecimento como faço esse link com, por exemplo, jogar, então, vamos supor que eu estou jogando isso aí lá para o LinkedIn. Eu não faço a menor ideia como que eu pego o que eu analisei e jogo lá para o LinkedIn, para isso aí ir passando para o pessoal.

Então, para mim, o desafio é esse, mas, tecnicamente, eu não sei nem dizer se é simples ou não.

Entrevistadora: Tranquilo. E a parte de teste e validação, assim, você consegue pensar algum teste para realizar, para ver se está funcionando bem, né, se está trazendo ali as pessoas que realmente têm fit com aquela vaga, e desenvolver, pensar em alguma métrica de sucesso, já que você está vendendo esse produto, né, que vocês estão lá sendo contratados.

E algum feedback para considerar as melhorias, as interações no algoritmo, né, pode ser feedback humano ou não humano.

Rebeca: Teste de validação seria quando entra algo novo no modelo, a gente faz o treino, né, tem o treino e teste no modelo, então, vamos supor que eu fiz um modelo de classificação, sim ou não, né, a pessoa estaria adequada para vir para a vaga ou não. Dado que aquele algoritmo jogou, por exemplo, dado que já está todo esse processo, né, jogou lá para o LinkedIn e me trouxe uma pessoa, eu vou fazer, tem lá, isso a gente faz com a mostragem, né, e aí eu separo em treino e teste, faço o treino, testo e aí vou ver se está funcionando, se essas variáveis que eu estou analisando e essas pessoas que estão selecionadas para ser no meu teste, está funcionando, está dando o resultado esperado.

Entrevistadora: Legal. E aí alguma métrica de sucesso, você consegue mensurar, para, tipo assim, você tem que entregar para o seu contratante, alguma quantitativa, né?

Rebeca: Não sei te dizer, Lari, assim, de imediato, mas, por exemplo, eu estava dando um exemplo lá do que a gente falou do Covid, né, que estava dando em quase 50

Entrevistadora: Boa, tá, entendi. E você considera ouvir as pessoas, por exemplo, as pessoas que estão na sua empresa, né, começar a trazer algum feedback ou na hora que você estiver fazendo mesmo a modelagem, que é um ambiente, assim, que você não é do RH, né? Então, você vai tentar dar o fit ali, com a vaga e tal, mas você considera ter essa interação com outras pessoas

enquanto estiver construindo o modelo, o algoritmo?

Rebeca: Com certeza, isso daí faz toda a parte desse entendimento que a gente está falando de trazer variáveis, né, porque, por exemplo, quando a gente falou ali, ah, eu traria ali, acho que, a formação, eu traria a experiência profissional, o RH, ele provavelmente já tem conhecimento muito maior do que de outras variáveis que pode ser analisada, que ele olha e fala, cara, essa pessoa está adequada ou não. Então, e não só com o RH, né, como que a área que também que está contratando, porque ali a gente está tendo, olhando ali, ah, a gente precisa disso, disso e disso.

Tá, mas para que exatamente? O que que vai executar? O que que vai fazer? Então, acho que é essa interação, tanto com a área que está solicitando, como com a área que tem experiência, que seria o RH, né, de contratação de pessoas, seria fundamental.

Entrevistadora: Legal. E aí, por último aqui, é mais uma relação, assim, de você entender, porque você vai fazer um algoritmo que vai selecionar pessoas e, a partir dali, essas pessoas vão receber a oportunidade, vão se candidatar ou não.

Então, já está havendo um recorte inicial, né? Você acredita que a maneira que você vai modelar esse algoritmo traz um impacto, não só, assim, na contratação específica, como na sociedade, uma coisa assim que pode trazer algum impacto mesmo, assim, naquele recorte?

Rebeca: Com certeza, porque, para começar, você já está... é que quando você falando de tecnologia, que está contratando para uma vaga de tecnologia e com experiência de cinco anos, você já está fazendo um recorte que você quer pessoas que tenham acesso à internet, à computação. Então, isso já é um recorte na sociedade, né? Você já está falando que você quer pessoas que tenham, entre aspas, uma classe social já... melhor, vamos dizer, né? Quando você faz um tipo de algoritmo desse jeito, você não está chegando lá de baixo e tentando pensar... não sei, Larissa, estou respondendo muita besteira, mas você não está pensando em, tipo, você não... falando em trazer uma melhoria para a sociedade, não necessariamente, eu acho. Socialmente, eu estou pensando, tá?

Entrevistadora: Está falando de melhoria ou de diferença?

Rebeca: Como assim?

Entrevistadora: Você falou assim, ah, trazer uma melhoria para a sociedade, foi isso que você falou.

Rebeca: É, trazer um equilíbrio na sociedade, né? Tentando trazer, assim, tipo, pensando que a gente vive uma desigualdade social muito grande, é um abismo, né? E aí, só que quando eu estou falando para uma vaga, que eu estou

já querendo trazer pessoas que tenham experiência com a tecnologia, eu já estou trazendo, eu já estou excluindo já uma parte da sociedade, é porque é complexo explicar isso, mas, por exemplo, eu já estou selecionando uma canata da sociedade, e aí eu estou mantendo ali, né? Eu estou com aquele meu miolinho, não sei se eu estou viajando, né?

Entrevistadora: Não, eu entendi, você já está fazendo, selecionando os requisitos que você está colocando, já está selecionando pessoas que já tiveram privilégios, né?

Rebeca: Exatamente. E aí você falou assim, ah, o que isso daqui pode melhorar para a sociedade? Cara, eu, porque quando eu penso em querer melhorar a sociedade, para mim, está diretamente relacionado a trazer equilíbrio social. E tudo, e cada vez mais, eu acho que a gente traz no nosso ambiente que a gente vive, mais desequilíbrio, mais desigualdade.

Então, ah, eu quero uma pessoa que tem experiência, porra, se eu quero uma pessoa que está com cinco anos de experiência, eu estou deixando aquela pessoa que está ali, às vezes, acabando de se formar, então, que não tem conhecimento, eu estou deixando aquela pessoa que não tem acesso à computação, já de lado. Então, por exemplo, vou dar um exemplo assim, o que eu estou querendo dizer? Eu, quando eu saí da faculdade, eu não tinha computador, e aí eu fui cair no mercado de trabalho, assim, eu tive a sorte que era, que estatística ainda estava sendo bem contratada, e ainda não era tanto o mundo como está hoje. Mas, hoje em dia, uma pessoa que se formou lá na UNESP, como eu, que não tinha computador, está muito atrás daquela galera que tem oportunidade, tem computador, tem tudo em casa, que é, né, e aí eu estou, quando eu estou criando uma vaga, ah, eu quero com experiência, seja esses cinco anos de experiência, não necessariamente pode ser cinco anos de experiência no profissional, ela pode ter um conhecimento que ela, estudando e lá, e fazendo, e durante a faculdade, sei lá, XPTO, e aí eu estou já causando, eu estou sempre selecionando a galera que é privilegiada.

Entrevistadora: Entendi. Legal, Rebeca, foi isso, muito obrigada, foi muito bom, é muito bom, é isso mesmo, assim, você foi, passou no teste, tá?

Rebeca: Passei no teste.

Entrevistadora: Foi tudo ótimo, eu vou te passar um link, logo agora, em seguida, do Forms, só para falar alguns dados específicos seus, que aí depois eu vou misturar com o que eu peguei, para fazer análise, mas já me ajudou muito, tá? Muito obrigada.

Rebeca: É que é um caso difícil esse lá, de pensar assim, de pensar na vaga, você fala assim, o que que é, e é uma coisa que você está explorando, que muita gente, acho que nunca pensou, porque sempre a gente está pensando,

ah, vamos falar da saúde, então eu vou lá, quando eu estou pensando, como a gente estava conversando no CIDEF, cara, abre um leque na minha cabeça, porque é coisa que eu já tenho pensado sobre, agora que eu falo assim, ah, pensa sobre divulgar uma vaga, é algo que é muito, e aí, quando vem o novo, você fica assim, cara, tem que pensar, tem que pensar sobre o que está fazendo sentido isso mesmo.

Entrevistadora: E falando de dados de pessoas, né, sempre muita, muita coisa, né, muitas vertentes que você tem para pensar, mas legal, Rebeca, muito obrigada, viu? Vamos continuar.

As duas: Tá bom, beijos, tchau, obrigada.

B.19 Regina

Entrevistadora: (...) que são perguntas mais abertas, tá? Não vou, a ideia, eu vou perguntar questões técnicas, mas eu não vou avaliar a resposta, a técnica da resposta, e sim, o que importa para mim é a construção dos seus caminhos de tomada de decisão, tá? Então, assim, é seu mapa mental para fazer, né, realizar aquela solução que eu vou passar. Então, tudo que você puder compartilhar de pensamento, caminho, você pode depois voltar atrás, dizer que não sabe, não tem problema, tá? É bem nesse caminho mesmo.

Regina: Tá bom, beleza.

Entrevistadora: Eu vou compartilhar aqui para a gente ir olhando, você está vendo a minha tela? Para a gente ir tendo como norteador mesmo, tá? Mas são os principais direcionadores aqui da nossa conversa. Então, queria, se você puder, primeiro, me contar um pouco da sua experiência com algoritmos, tá? O que você desenvolve hoje em dia também na sua empresa, não precisa entrar em detalhes de cliente, nem de nenhum projeto, mas é mais em relação à sua experiência mesmo, pode ser profissional, pode ser acadêmica, você achar que vale aqui no contexto.

Regina: Minha experiência no que eu trabalho hoje em dia é bem diferente do que eu estudei, assim, eu fiz engenharia química e acabei que eu não entrei na área, acabei me formando numa época que não tinha muito espaço para engenheiro, então me guiei para a área de dados.

O que se repete muito, realmente, ali é a hora de construir a lógica, né? Ver, entender primeiro, eu quero o que o meu cliente espera, né? É uma coisa que nesse tempo eu tive que aprender, né? Com o tempo na empresa. Você mostrar para ele o que que ele quer ver, se faz sentido ou não, é uma coisa que é complicada, tem que estudar muito o que que você tem, que não adianta você achar que tá ali fácil, tá na cara, mas não tá, isso é importante. Então,

o meu papel, eu acho que hoje em dia é conseguir mostrar mais o que que a gente tem e tentar mostrar, acho que, possibilidades, né? Do do que podemos chegar, acho que tô tô nesse caminho.

Entrevistadora: Tá legal. É, vou te passar uma contextualização de um cenário, tá? Que eu criei. Você imagina, assim, uma empresa de uma tecnologia, desculpa, uma empresa de tecnologia de porte médio, tá com vaga aberta para contratar pessoas, e aí ela precisa, quer divulgar essa vaga, essa oportunidade, só que ela quer divulgar, ela quer que essa vaga chegue só para as pessoas que já tenham algum filtro ali, né? Já tenham algum direcionamento que possam dar médicos essa vaga.

Então, ela vai contratar uma outra empresa para fazer a divulgação dessa vaga e vai passar os requisitos e pedir para desenvolver um algoritmo, alguma coisa, né? Desse tipo, para já selecionar essa vaga, essa oportunidade, vai aparecer para as pessoas específicas, assim, por exemplo. Não vai aparecer para pessoas que são formadas em história, vai aparecer para pessoas que já tenham alguma formação na área técnica, né? Uma coisa mais para os requisitos que eu vou falar aqui. Então, o que que eles passaram? A vaga de tecnologia para um profissional pleno, com skills que contém um desenvolvedor ou uma desenvolvedora full stack, formação em engenharia de software, tecnologia da informação, sistemas e áreas afins, experiência em média de cinco anos e para trabalhar em regime remoto.

Tá? Então, assim, sabe esses anúncios que aparecem no Instagram? Sim. Que já tem alguma coisa que já leram ali de você, que você pode se interessar por aquele anúncio. Então, a ideia é isso, que tenha um algoritmo que vai já direcionar para as pessoas que podem caber nessa vaga.

Então, você, que é a pessoa que trabalha nessa empresa, que vai fazer o algoritmo, eu queria que você me dissesse como que você planeja realizar, né, dar conta dessa situação que eu passei, dessa, como é que você conseguiria desenvolver essa solução? Então, no primeiro momento, em relação ao planejamento e concepção, como você faria o refinamento, né, definir melhor esses requisitos que já foram passados ali, mas como você traduz essa definição de requisitos no pensamento de desenvolvimento de sistemas e quais dados utilizados, quais dados você acha que você deveria puxar para responder aqueles requisitos que foram elencados ali em cima? Vou voltar na primeira página, e aí ficam com os requisitos aqui. Vaga de tecnologia para profissional plenos, essas aqui, né, tudo isso que eu li aqui.

Regina: Volta lá, nas perguntas, os requisitos, tipo de dados.

Então, teria que primeiro ver... volta lá, por favor. Assim, o tempo de empresa, por exemplo, é um pouco diferente na hora de ver os cargos, então

talvez tem que dar uma atenção maior assim para o tempo de cargo, na hora de fazer uma avaliação. Porque, por exemplo, a gente pega ali o LinkedIn, então tem várias possibilidades.

Acho que primeiro a gente tem que ver o que é mais importante ter preenchido. Então, na hora de pegar ali a base, você vê, por exemplo, eu imagino que a formação tem que estar, a maioria, preenchida, mas às vezes não está suficiente. Então, você tem que, talvez, nesse caso, acho que criar uma pontuação.

Um cenário, como você explicou aí, eu acho que talvez criar uma pontuação na hora de organizar as pessoas, para selecionar o melhor. Tá, mas aí a pontuação relacionada a quê? Você olharia a formação? Você falou aí da empresa, assim, da experiência na empresa? Eu acho que seria primeiro o tempo de experiência, talvez. Ver se a empresa que ele trabalhava tem a ver com uma área, né? As habilidades, porque, né, o que você já aprendeu, porque é, vou falar aqui, tem formação disponível, né? E é tudo muito curto, de certa forma.

Você consegue agregar muita informação, então você detalha uma quantidade de curso maior, talvez seja uma uma coisa de peso, né? Uma característica de peso. A formação, acho que formação tem que começar, assim, do que tem mais a ver, para o que tem menos, talvez, a ver, né? Que nem engenharia, fazer engenharia, que nem eu, versus outra pessoa que fez história. É bem diferente, né? E regime remoto, acho que é mais fácil de pessoal aceitar, então acho que não tem muita importância, né? E acho que seria assim.

Entrevistadora: Tá. Então, você sabe dizer qual método você escolheria? Se você escolheria uma técnica de machine learning, ou não precisa ser, né, mas outra técnica, para desenvolver esse algoritmo? E quais desafios técnicos você acha que você poderia tentar? E aí, talvez isso afete, inclusive, a escolha do seu método. E esses dois são os principais, né? E aí é relação à integração, implementação, isso se você imaginar que você vai fazer, por exemplo, um algoritmo que vai integrar com uma plataforma tipo LinkedIn, mas é mais assim, em relação ao método e os desafios mesmo.

Regina: Seria, né, machine learning, seria uma coisa, um modelo mais simples, né? Acho que, dependendo da forma que for avaliar, acho que seria mais fácil de, deixa eu pensar, bom, você tem ali os dados das pessoas, você quer acertar qual que é o melhor candidato, né? É isso que você quer.

Entrevistadora: Isso.

Regina: Bom, acho que machine learning realmente é isso, acho que o desafio seria saber a confiabilidade do dado, né? Como que eu não sei de onde que vem, talvez trabalhar os dados, melhorar a qualidade deles, não tem muita

dificuldade nessa hora.

Eu acho que seria uma coisa não muito difícil de acertar, mas tem que saber informações que estejam preenchidas, né? Evitar valores vazios. E na hora de pensar na integração, hoje em dia tem muita possibilidade de conectar facilmente ou rodar facilmente um programa, então não seria muito difícil, né? Dependendo de como saem os arquivos e como é feito o processamento, não é muito complicado.

Entrevistadora: Legal.

E sobre os testes, depois que você rodar esse algoritmo, aí ele vai selecionar as pessoas, né? E as pessoas que vão receber esse anúncio, você, como você realiza esses testes para ir aferindo o modelo, né? Você, como você também, se você pensa em alguma métrica de sucesso, até porque pensar que isso é um produto que você está desenvolvendo para uma outra empresa que contratou, né? Então, você tem que dar algum feedback ali. E quais feedbacks você consideraria para a iteração, né? Para ir melhorando o algoritmo. Você considera feedback interno, externo, humano, não humano, se você consegue também dizer sobre isso?

Regina: O teste seria, tem na hora de fazer o machine learning, ele já faz alguns testes, né? Retira amostras, vê se está acertando.

Então, dá para repetir, teria que, pode testar com os resultados obtidos, né? Uma forma de avaliar se está ficando bom é, ao longo do tempo, ver quanto tempo a pessoa está na empresa. Se ela continuou, né? E seria bom saber algumas características que talvez estejam de acordo, né? Então, a empresa contratante esperava que ela conhecesse algumas informações, ela poderia dar um feedback para a gente, verificando ou não, talvez se essas pessoas estão correspondendo com as características, né? Que eles queriam. Seria, acho que, uma coisa mais no começo, né? Se for para imaginar que está valendo direito o candidato.

E a questão de feedback é, talvez, até o funcionário, né? Ou a pessoa que foi contratada, seria legal saber se atendeu a expectativa dele. Talvez, se ele também procura uma empresa, a empresa também tem que achar ele. Dependendo como que foi indo, ao longo do tempo, daria para até fazer uma, um paralelo, né? E o feedback é isso, acho que, ver se todo mundo está atendendo a expectativa, tanto interno quanto externo, depende do que a pessoa está te falando, né? Que nem a parte técnica, uma coisa mais interna, enquanto que falar acertou, não acertou, é o cliente, né?

Entrevistadora: Sim.

Você considera isso importante quando você vai desenvolver um algoritmo, vai pensar nos pesos, né? Nas notas ali que você vai dar, você costuma

conversar com seus pares sobre isso, para ter uma decisão mais discutida, e você acha importante ter esses feedbacks do cliente, né? Para ir ajustando também?

Regina: Com certeza. O responsável pelo negócio somos nós, então tem que ter a participação deles. Pode ser que eles já tenham a informação já feita em algum local, mas eles têm que participar, se daí eu concordar de algum material que eles já possuam, né? A gente vai lá, talvez, se tiver essa sugestão, né? Eu pegar e falar, eu acho que essa ordem tá boa, você concorda? Se eles derem um ok, tudo bem, né? Mas tem que, o ideal seria conversar.

Acho que todo ali registrado, via e-mail, pelo Teams não é o ideal, por exemplo. Tem que sempre ter a participação da empresa.

Entrevistadora: Legal, tá bom.

E em relação aí a uma opinião sua, tá? Você acha que, em relação à efetividade, você acha que isso funciona bem? Fazer esse algoritmo, deixar um pouco mais automatizado o anúncio dessas vagas, para buscar mais pessoas, né? E atingir mais possíveis candidatos, se isso acelera ali ou não, se isso ajuda na efetividade realmente da contratação. E em relação, quando a gente tá falando de algoritmo, né? A gente tá falando de vaga de trabalho, esse algoritmo vai escolher quem são as pessoas que vão ver esse anúncio, né? Isso tem algum impacto, inclusive, social, né? Você considera que é importante, tem alguma questão mais sensível, em relação a isso? O que você acha?

Regina: Bom, impacto social, que pergunta complicada. Bom, acho que entra no mérito de como que tá avaliando, né? Quanta informação ele tá puxando da pessoa para fazer essa constatação, né? Por exemplo, se for alguma coisa que tá nas redes sociais, hoje em dia a gente não consegue controlar o quanto que eles estão recebendo de informação nossa, aqueles cookies, né? A gente só aceita.

Então, dependendo do quanto o algoritmo tá recebendo de informação, pode dar uma tendência ali no produto, né? E a efetividade, acho que volta na questão de ver se as pessoas estão encontrando o que elas procuram, né? A recomendação de quantidade de dados com as skills desejadas. Eu acho que realmente é a parte de avaliar, talvez. E sempre comparar os resultados, né? Que nem eu falei, ah, você tem os testes, você tem que saber até onde que é o seu interesse que acerte, né? Pode ser que seja baixo ou você quer o mais alto, né? Aí a gente vai vendo, talvez, alguma forma de incentivar o candidato a preencher alguma informação.

Por exemplo, dá para forçar a pessoa a colocar o, sei lá, onde que ela estudou, vamos falar assim, quer algo mínimo que vai ter. Então, poderia ser uma possibilidade, caso a pessoa esteja se candidatando pelo formulário, uma

vaga já lá no Google.

Entrevistadora: E deixa eu te perguntar, última pergunta.

Você acha que, nesse caso, né? Como eu falei do impacto social, é porque assim, você vai escolher quem vai receber essa oportunidade. Então, nesse sentido, se o seu algoritmo ficar enviesado, ele pode escolher determinados grupos e despriorizar outros. Você acha que, nesse caso, que existe inclusive um impacto direto assim, da organização social, enfim, você acha que teria, seria importante, se tivesse um checklist na hora de você fazer esse algoritmo, com algumas questões assim, por exemplo, para te direcionar no desenvolvimento mais inclusivo, mais igualitário, enfim, com questões que talvez não passem pela sua realidade, que você não vai pensar, mas que traz ali a reflexão para você ser obrigada, vamos dizer assim, né? Você ser incentivada a pensar em outras situações para tentar, na hora de desenvolver o algoritmo, de imaginar situações, você ser mais empática com outras questões, outros grupos, de repente, né? Você acha que seria interessante isso?

Regina: Ah, legal, acho que uma possibilidade ao longo do tempo, né?

A gente primeiro tenta acertar essa parte de conseguir estar acertando a informação, acho que tem que pensar na entrega, né? Então, a gente primeiro atende esse pedaço, mas pensar para frente, entender como que, que informação é importante nessa hora de avaliar se o candidato pertence ou não ao grupo que a gente pretende atingir, né? Se for um grupo mais amplo, como pessoas pretas, mulher, PCD, você começa ali a colocar de variável no processo, né?

Antes você poderia não estar considerando, mas hoje em dia você considera na hora de obrigar a pessoa a responder, né? Que eu tinha dito anteriormente.

Então, não é complicado, assim, na hora de você pensar no código. Por exemplo, eu acho que ele também vai entrar na hora de avaliar o quão importante é, tem que também pensar nisso. Então, na hora de programar novamente, vai ter sua importância, mas não é igual à experiência, né? Tem coisas que são com um peso bem maior, né, na hora de avaliar, mas dando o devido valor à informação.

Então, seria por etapas. Creio que dá para atender e realmente é importante, faz diferença na hora da empresa estar aberto ao público, né? É meio complicado algumas perguntas que eles fazem, né? Tem empresa até que pergunta, ah, você casou faz quanto tempo? Você pretende ter filho? Tipo, tem coisas que não faz sentido, mas acontece. Então, tem que estar aberto.

Entrevistadora: Muito bom, Regina, era isso. Muito obrigada, viu? Vai me ajudar bastante. Obrigada pelo seu tempo.

Regina: Obrigada. Obrigada, eu. Beijo, até mais.

Entrevistadora: Tchau, tchau.

Entrevistadora: Norteadores, tá? E aí a gente vai usar também o método que a gente usa em design, chama Directed Storytelling, que é você, eu passo alguma ação, né? Você vai fazer uma, vai desenvolver uma ação, e aí você vai me falando tudo que vai vindo à sua cabeça, quais as decisões você tomaria, né? É bem assim, focado na ação, né? E aí você vai narrando. Eu vou compartilhar aqui, mas assim, se você quiser tomar a tela para qualquer coisa, mostrar alguma ferramenta, algum, sei lá, uma plataforma que você queira dividir para ilustrar melhor o que você tá explicando, fica à vontade, tá? Você já tá vendo a tela? A princípio eu gostaria que você contextualizasse melhor a sua experiência com algoritmos, pode ser profissional, acadêmica, né? Qualquer coisa que você tenha tido contato, e falar um pouco do seu papel na empresa hoje, que você vem trabalhando também.

Renata: A minha experiência com algoritmos, ela começou com o desenvolvimento front-end, então eu trabalhava com JavaScript, HTML, CSS, foi o início da carreira.

Depois eu fui trabalhar com back-end, então eu já mudei, né? Eu já fui trabalhar com Java, já fui trabalhar com Hibernate, com frameworks Java. SQL eu trabalhei com Oracle, com MySQL, com SQL Server também, o outro tem quatro. MySQL, SQL Server, Oracle, ah, e Postgres, pronto, lembrei dos quatro.

Aí trabalhei uns anos com isso, para depois começar a trabalhar com Python, então assim, foi uma migração. Quando eu trabalhava com Java, era para desenvolvimento de aplicação. Agora, quando eu mudei para trabalhar com Python, então já foi começo com desenvolvimento de script, e aí foi a migração para dados.

Eu já comecei a trabalhar com outro tipo de algoritmo, não era para uma aplicação exatamente de OLTP, mas agora começar a trabalhar especificamente com o script para manipulação de dados. SQL eu sempre usei em todos os anos de carreira, mas atualmente eu trabalho mais com Python, porque eu manipulo dados, que torna um pouco mais fácil esse trabalho. Com Java, ele é uma linguagem bem mais verbosa, e bem mais difícil de trabalhar com isso, e dá toda essa facilidade, porque a linguagem de Python é de script, então fica muito mais simples.

Atualmente, na empresa, sou consultora sênior, então eu atendo clientes, então eu faço algumas atividades técnicas com Python também, para algoritmos, especialmente para manipulação de dados, porque eu trabalho na área de analíticas, mas eu tenho também esse viés de conversar com o cliente, então não é só algoritmo. Até na definição do algoritmo, eu tento entender muito bem qual é o problema que está sendo descrito, para modelar ele de uma forma

melhor.

Entrevistadora: Muito bom, legal.

Eu vou então começar, vou explicar um pouco do cenário do objetivo dessa simulação, vamos chamar assim. Imagina que uma empresa de tecnologia, de porte médio, contrata uma empresa que você trabalha, né? Então, legal que você já trouxe aí, vamos dizer, e todo esse background aí de consultora, né? Exatamente esse papel que você teria dentro dessa empresa. E aí, então, essa empresa de tecnologia de porte médio, contrata uma outra empresa para fazer o anúncio, rodar um anúncio de uma oportunidade de trabalho, tá? E aí essa oportunidade, eles dão alguns requisitos, que é uma vaga de tecnologia para profissionais pleno, com skills de desenvolvedora ou desenvolvedora full stack, formação, engenharia de software, tecnologia da informação, sistemas e afins, uma experiência média de cinco anos para trabalhar em regime remoto.

Só destrinchando assim, sabe esses anúncios que aparecem no Instagram, Facebook, né? Assim, que você sabe que já teve algum recorte ali de perfil para aparecer para você especificamente, né? Então, o algoritmo já entendeu que você pode se interessar para aquele anúncio. A ideia é exatamente isso, fazer um algoritmo para ler o perfil de algumas pessoas, né? O perfil profissional de algumas pessoas, para que elas, assim, né? Fazer uma estimativa ali que aquelas pessoas vão dar match, vão dar fit ali com a vaga que está sendo anunciada, com esses requisitos, tá? Não precisa se ater, assim, a questões técnicas de plataforma, né? Assim, ah, não sei se dá para fazer no LinkedIn e tal, vamos colocar no mundo ideal aqui, tá tudo liberado, todos os dados, enfim, pode pensar nesse cenário, tá? Aí, a primeira etapa é o planejamento e a concepção. Então, tem a definição de requisitos, né? Como você refina, vamos refinar, porque tem alguns requisitos já foram passados ali para o algoritmo, né? Para inserir isso no algoritmo e quais os dados você utilizaria, né? Em relação para conseguir contemplar aqueles requisitos específicos.

Eu volto aqui para você ver.

Renata: É, isso eu queria ver, pronto. Eu acho que dá para trabalhar com palavra-chave, acho que é melhor para poder aparecer para as pessoas, né? Pensando nisso, né? Então, seria interessante o algoritmo conseguir se ler as descrições que as pessoas colocam em qualquer plataforma, qualquer uma, anyway.

Então, buscar essas palavras-chave no perfil da pessoa. Por exemplo, para um profissional pleno, talvez olhar quantas experiências ela tem e o tempo médio, né? Que ela passou ali. Cinco anos, nossa, a mente está cravada em cinco anos.

Eu não sei se foi uma exigência, né? Mas, olhar mais ou menos quantas

experiências ela tem ali e qual o tempo médio que ela passou ali, o algoritmo consegue calcular qual foi o tempo da pessoa. Como a desenvolvedora do Fustec, seria interessante que nessas palavras-chave tivessem umas palavras importantes dali. Github, Git, a palavra em si, desenvolvimento do Full-stack ou não.

Mas, alguma linguagem de programação seria interessante estar mencionada ali, alguma linguagem. Alguma palavra-chave relacionada à metodologia ágil, já que ele é uma pessoa plena, então, provavelmente, ela já teve outras experiências. Seria interessante que ela tivesse também.

Outra coisa que eu olharia na faculdade dela, olhar as faculdades que têm alguma coisa relacionada com software, informação, computação, sistemas, que são os nomes mais usados, construídos para fazer essas informações de tecnologia. Eu sou mesmo de sistema de informação. Então, sistema de informação, ciência da computação, análise e desenvolvimento de sistemas, engenharia de software.

Outra coisa importante, já que é um profissional pleno, ele tem que saber algum nível de banco de dados, então, alguma menção a algum banco de dados. Tanto o banco de dados de SQL, que é um banco relacional, e até um banco não relacional. Seria interessante, é um profissional pleno.

Deixa eu ver o que mais. O regime remoto. Em geral, todo mundo está escolhendo o remoto, né? Então, assim, é difícil ver um profissional colocar no currículo.

Ah, quero somente presencial. Eu até hoje não vi, recentemente eu não vi. Eu já vi híbrido e remoto, vocês colocando como preferência.

Mas, se tiver opção, vai presencial mesmo. Então, eu não vejo. É importante a empresa divulgar, porque isso também chama a atenção do profissional, na hora que o anúncio, porque é interessante que o anúncio apareça, que você interage com o anúncio.

Então, eu colocaria o regime remoto. Mas, para aparecer, eu acho que tentar pegar as palavras chaves do currículo da pessoa. Não ia levar em conta nada de nome, nem ser algo que remeta ao gênero da pessoa.

Eu acho isso irrelevante. Então, principalmente o texto currículo, que ela fala dela mesma, também. Esse é um texto que descreve as habilidades.

Então, procure essas palavras chaves que descrevem os requisitos que eu preciso, como profissional.

Entrevistadora: Legal. Daí, assim, em relação ao desenvolvimento do algoritmo, você entendeu melhor ali os cenários, requisitos e já começou a definir os dados que você vai direcionar, qual método que você escolheria, e pode ser alguma técnica de machine learning ou não, tá focado nisso.

E, até assim, se você consegue mensurar algum desafio técnico, que tenha, assim, até influenciar na escolha do seu método específico. E essa parte de integração e implementação, se você tiver alguma ideia, mas não é fechado, tá?

Renata: Como eu estou trabalhando com a questão de texto, eu acho que, na verdade, seria interessante fazer uma técnica híbrida, porque nesse tratamento do tempo de experiência, então, de alguma forma, você teria que capturar quais são as experiências da pessoa e o tempo médio. Essas plataformas sempre têm, qualquer plataforma, né? Você coloca o início e o fim.

Então, poderia usar alguma feature engineer para pegar essas features relacionadas a tempo de, não sei qual é o nome, o tempo de trabalho naquela empresa, naquela organização, para poder fazer o tempo médio de experiência da pessoa. Então, teria que fazer alguma coisa um pouco híbrida. Features engineers relacionadas a capturar esse tempo e algumas técnicas de NLP para processamento de texto.

E aí, você pode transformar esse texto em variáveis categóricas, para poder usar no modelo de classificação, de ranqueamento, alguma coisa assim. Mas, para mostrar fulano é o potencial. Então, apresentar isso.

Por quê? Para capturar as palavras-chave. Então, construções também. Dá para usar um n-gram para poder capturar as construções de frases.

Porque, às vezes, não é uma palavra, é uma expressão que é mais utilizada. Então, também usaria isso para processar o texto. E seria uma multi... deixa eu ver se recomendar ou não recomendar.

Então, dependendo de onde apareceu o anúncio da pessoa, dá para trabalhar com técnicas de recomendação também. Eu acho que a classificação é uma recomendação. Depende de como eu gostaria de apresentar isso.

Quais seriam os principais desafios técnicos enfrentados durante o desenvolvimento do algoritmo? Deixa eu pensar. O processamento dos dados seria um pouco complexo, porque depende da plataforma. Se eu vou ler um currículo PDF, se eu vou ler um currículo Word, eu poderia copiar no currículo, na leitura do currículo da pessoa, né? Então, depende como essa plataforma iria trabalhar com os meus dados.

Se ela for uma plataforma que consegue pegar os campos do currículo, fica um pouco mais fácil, porque eu consigo tornar isso, assim, mais palpável, organizar isso no banco de dados. Mas, se não, tenho que usar uma outra tecnologia que possa ler esses currículos. Então, o formulário é muito mais fácil, né? Eu pego todos os dados da pessoa, tenho lá bonitinho o modelo.

Mas, se na verdade eu não tenho esse recurso, só tenho um currículo

PDF da pessoa, então eu vou ter que ter alguma forma de ler aquele PDF, ler aquele documento Word, ou qualquer formato, para poder capturar as minhas informações. Eu acho que seria uma das principais questões. Como o algoritmo foi integrado no sistema de divulgação de bagagem da empresa? Eu acho que, se for fazer uma plataforma com o LinkedIn, eu não sei como funcionam os bastidores para você, com aquele trabalho lá internamente.

Mas, pensando em sistemas de propaganda, alguma coisa assim, eu acho que posso usar os desafios é encontrar onde eu posso anunciar para que a pessoa veja. Por isso que a gente fala tanto do LinkedIn, porque todo mundo entra. Mas, talvez, seria importante anunciar em outras plataformas.

Então, acho que a grande questão é como funcionam os bastidores dessas plataformas para poder recomendar para a pessoa.

Entrevistadora: Muito bom. E aí, assim, pensando que você foi contratada, né? Então, está desenvolvendo um produto, né? Você consegue mensurar quais testes você realizaria para ver, né? Assim, para validar o algoritmo.

Tem alguma métrica de sucesso, alguma questão quantitativa ou qualitativa mesmo, que você consegue mensurar para retornar isso para a empresa que te contratou? E você considera feedbacks para fazer iterações humanos, não humanos? Você acha importante?

Renata: Olha, eu faria um experimento interessante com o teste A-B, no tipo de usar uma abordagem ou outra e ver qual que eu tive mais sucesso, mas de forma duplicada mesmo, não vai saber a pessoa, que técnica que está sendo utilizada nela. Tentaria fazer isso, ou tentaria também fazer sem técnica nenhuma, e eu oferecendo. Então, fazer algum tipo de teste A-B, para ver qual seria a real diferença, né? De recomendar para ela ou não recomendar.

Eu estava pensando aqui também, também fazer segmentações, né? E por plataforma. Então, nessa plataforma eu faço desse jeito, nessa outra. Qual que no final, aí eu tenho que ver com a minha métrica de sucesso.

É a pessoa se inscrever na vaga? A pessoa olhar a vaga? Isso seria já uma questão que eu iria definir com a equipe de produto. O que que para vocês é um número relevante? É número de visualizações? Entre visualização e quem se inscreveu? O que que seria interessante, né? Como para avaliar? Para um negócio, eu acho que eu iria primeiro ver isso. O que que é uma métrica de sucesso para o negócio? Porque existe aquela coisa de desempenho de modelo, é uma coisa, né? Precisa um recall, a curar, assim, anyway.

E outra, o que que é uma métrica para o negócio? Que eu acho que, às vezes, isso é um pouco deixado de lado. E aí, por isso, alguns modelos entram no limbo. Mas, para o negócio, o que que é importante? É que as pessoas vejam, que as pessoas se inscrevam, as interações.

Então, isso eu definiria com a equipe de negócio, produto, seja quem for. Feedback, como avaliar feedbacks interno ou externo, que é a modificação do algoritmo. Vamos lá.

Exatamente essa interação com o pessoal de negócio. O que que é importante para vocês? Então, eu iria criar essa métrica, que é uma métrica customizável. Claro, tem a do algoritmo, mas eu acho que é importante olhar para o negócio e criar a própria métrica para validar se está sendo efetivamente eficiente para o que se propõe.

Então, eu iria falar com o pessoal de negócio, vamos criar uma métrica, primeiro. Essa métrica vai ser essa. E aí, a partir dos resultados, a gente vai conversar e tentar entender.

Está adiantando? Está funcionando para vocês? Está trazendo o resultado que vocês esperam? E aí, a depender, a gente vai fazendo algumas modificações. Claro, eu vou avaliar as questões mais de... O cliente vai dizer para mim. E eu, tecnicamente, como eu conseguiria atuar para resolver isso? Vou mexer no modelo? Vou mexer nos dados? Vou mexer na plataforma? Então, dependeria muito, assim, das conversas, né? Porque ele me fala a demanda, esses testes de saber vão responder um pouco disso, mas a questão técnica eu vou ver onde é que eu tenho que atuar.

Entrevistadora: E você considera também, assim, na hora da construção mesmo, do algoritmo, quando começa a fazer a montagem do modelo, você considera conversar com pares, com pessoas da sua equipe, você acha isso importante? Ou você... Você falou bastante em falar com o pessoal do produto, né? Para definir as regras de negócio e tal, refinar ali. Mas você, até pode ser do seu dia a dia, né? E transpondo para essa situação, você acha importante ter discussões internas, assim?

Renata: Acho, e eu já faço isso. Para mim, por default, tem que ser feito.

Ninguém pode tomar decisão unilateral. Porque a gente tem um problema de viés grande. Quando a gente só consegue enxergar um lado, então talvez a outra pessoa que está do teu lado está enxergando uma outra questão, um outro problema, uma outra perspectiva.

Então, para mim, é um default, tem que ser feito isso. A equipe tem que ser conversada. Como que a gente vai fazer isso? O que você notou, principalmente para variar o que o cliente diz? Porque como você capta a mensagem é diferente de cada um.

Então, eu posso falar o que eu entendi. Ah, eu entendi isso aqui. Ah, eu já entendi dessa forma.

Então, isso ajuda você a ter uma visão mais ampla do que foi dito, do que foi pedido.

Entrevistadora: Ah, legal. E aí, assim, em relação... Como a gente está falando, né? De algoritmos que vão selecionar um certo grupo para oferecer uma oportunidade, né? Você, em relação à efetividade, você acredita que é efetivo fazer essa escolha, né? Botar um algoritmo para fazer a seleção de pessoas.

E em relação ao impacto na contratação, como a gente está falando, né? A gente está falando de pessoas. Você acha que o algoritmo, ele tem um poder de impacto grande em fazer essa seleção e resultar em contratações. E pensando até em uma questão mais de sociedade mesmo, né? Como que o algoritmo pode impactar nesse direcionamento dessas pessoas e fazendo esse recorte de oportunidades.

Como que você enxerga isso?

Renata: Eu acho que tem dois lados. Vamos lá. Pensando no RH, é muito positivo que ele vai ganhar tempo.

Então, ao invés de você estar procurando as pessoas, ele acha para você. Então, isso tem um ganho de eficiência para o pessoal de recursos humanos da empresa. Para avaliar se realmente está sendo efetivo para a empresa, então, seria interessante, após a contratação, avaliar como foi o desempenho dessa pessoa, né? Durante algum X tempo que a empresa considera relevante.

Agora, olhando para o candidato. Tem um lado ruim também, porque você está participando de uma segmentação. Então, depende muito do que você coloca no seu currículo.

Inclusive, isso de alguma forma já é feito, né? Então, isso acaba tirando algumas pessoas. Isso é um fato. E depende de como esse algoritmo foi construído.

Então, você tem um problema de viés, sim. Ele pode selecionar um grupo. Mas é por isso que eu não ia colocar como informação pessoal nada relevante.

Nada disso, para que a gente possa diminuir ao mínimo possível os viés. E que tudo seja muito mais focado nas habilidades, no que você descreve como profissional no currículo. Porque esses algoritmos têm um pequeno probleminha.

Se você colocar coisas relacionadas, coisas pessoais. O gênero da pessoa, a data de nascimento, às vezes nos currículos vem. A região dela, que eu acho que é um remoto, isso não pode existir.

Não tem porquê eu contratar uma pessoa remota e exigir que ela more em um estado. Não faz o mínimo sentido isso. Então, não é remoto.

Então, eu tentaria diminuir ao máximo esse tipo, para que... Eliminar completamente o viés, você não vai conseguir. Mas, pelo menos que seja focado nas skills. E não no que a pessoa é. Se é mulher, se é homem.

Qual é a questão... Se é trans, se não é. Isso não é problema para mim, entendeu? Isso é uma coisa completamente irrelevante. Eu acho que relevância é a habilidade da pessoa. E outra questão é essa regional, que é complicada.

Mesmo remoto, depende de onde você anuncia. Como. Você pode, na verdade, estar criando um nicho ali, um bioma, uma região.

Então, eu tentaria eliminar um pouco isso. É remoto. O que você tem que ter é a infraestrutura para fazer isso.

Então, eu avalio esses dois lados. Mas, no final, eu ia analisar quem são as pessoas que o meu algoritmo está selecionando. Das que foram contratadas.

Claro, eu posso fazer dos currículos uma triagem, entender o perfil. Até para ver se realmente eu consegui eliminar os vies. E até se tem uma plataforma que crie um vies.

Porque, às vezes, aquela plataforma é mais utilizada por um determinado segmento de pessoas, uma classe social específica. Então, eu tentaria analisar isso. Quem ele está mostrando para mim.

Mas isso seria uma pesquisa interna. Eu não iria usar isso para alimentar o algoritmo. É só para eu entender e analisar como é que ele está performando.

Para poder fazer as modificações necessárias. Para detectar que a coisa está sendo injusta. Como o algoritmo pode gerar algum impacto significativo no processo de contratação da empresa.

Sim, é como eu disse. Tem um RH. E vai ter um ganho de eficiência.

Isso não é uma dúvida. Mas eu acho que é bom verificar depois como é que está sendo o desempenho dessas pessoas dentro da empresa. E até analisar como... Pegar o formulário de feedback das pessoas que são contratadas.

Para entender como foi o seu processo. O que você achou. Também analisar as plataformas para ver.

Tem tudo isso.

Entrevistadora: É isso. Nossa.

Vou te contratar, hein? Se eu fosse dessa empresa eu ia ficar feliz de contratar. Muito bom.

Renata: Que bom, que bom.

Entrevistadora: Foi ótimo.

Renata: Você está analisando na lista como é que os homens... Como os gêneros, na verdade, respondem essas perguntas.

Entrevistadora: Na verdade, assim... Pela minha pesquisa já está comprovado que existe produção de vies de gênero em algoritmos. Tenha até a inspiração para pegar essa... Deixa eu parar de gravar. Senão vai ficar muito grande.

Renata: Muito me interessa por esse tema. Depois que você tiver...

C

Transcrição: Grupo de Foco

C.1

Grupo De Foco (homens)- Checklist Para Desenvolvimento De Algoritmos Sensíveis

Entrevistadora: Oi, pessoal. O que que vocês acharam, vocês entenderam, acharam complexo, acharam tranquilo? O que que vocês entenderam, assim, a priori, né, sobre esse item? Quem quiser falar, pode, todo mundo tímido.

Davi: Acho que é um pouco, considerações um pouco diferentes do que estou acostumado a ter questionado, ou ser pedido no dia-a-dia. Não são coisas que geralmente são, estão um pouco fora da realidade, do meu dia-a-dia, esse tipo de ponderação.

Entrevistadora: E aí, você acha que, estando fora, é uma coisa mais, assim, pouco aplicável, ou poderia, pode entrar, assim, né, se você elaborar, claro, né, daria para entrar na sua rotina, assim?

Davi: É, eu acho que esse, a primeira, a primeira, a primeira parte, né, do que é pedido, eu acho que dificilmente, se eu entendi, eu não, para ser sincero, não entendi muito bem. A segunda parte, sim. A maneira como as pessoas mudam ao longo do tempo. Isso, sim, pode ser incorporado.

Entrevistadora: Mas o que que você entendeu sobre isso?

Davi: Bom, eu pensei do meu ponto de vista bastante prático, né, como, no caso de validação de identidade, a gente recebe imagens, por exemplo, de documentos, né.

Nesse caso, a gente trabalha com foto de um documento que pode ter sido tirada de 40 anos atrás. Então, a pessoa mudou ao longo do tempo. Então, isso é algo que a gente tem que levar em consideração para, na hora dessa análise, por exemplo, para validar se uma pessoa é aquela mesmo, né.

Ter exemplos de, como esse, né, de pessoas num espaço de longo tempo, espaço de tempo mais curto, né. Considerar essa mudança, essa possível mudança, a estética mesmo da coisa. Foi bem literal, a minha interpretação foi bastante literal em relação a isso.

Entrevistadora: Você, Bruno?

Bruno: Não sei se eu interpretei da maneira certa, não sei se existe certo, mas... É. O que que eu entendi, assim, no geral sobre esse primeiro card aí, no geral das coisas? Tipo, o que é que eu coloquei aqui? Tipo, as pessoas mudam de acordo com vários fatores, é individual de cada um disso, assim. Pode ser influenciado, sei lá, uma pessoa ansiosa, pode às vezes se dar mal, por exemplo, no desenvolvimento de alguma coisa. Às vezes pode ter ficado nervoso, não consegue tocar uma tarefa.

Uma pessoa ficar nervosa pode ficar, sei lá, pode ser uma pessoa muito boa tecnicamente e participar de uma entrevista e ir muito mal. E como que considera isso ao longo do tempo é que as pessoas amadurecem. Tanto por experiência profissional, às vezes por evolução não só de coisas técnicas, profissionais, mas tipo, a pessoa em si, né.

Então isso, tipo, muda e vai ter que ser adaptado, né. Em relação a isso de como seria realizado no dia-a-dia, eu acho que sim, é importante, mas dia-a-dia eu achei até um pouco muito um pace muito rápido. Tipo, isso vai adaptando conforme o tempo, observando cada um dos seus pares ali, o time que você estiver interagindo, as pessoas que você vai estar tendo essa interação.

Então, é recalibrável. Eu penso que sim, eu considero que é aplicável.

Como que você faria pra aplicar isso, né? Quais os meios buscaria? Eu tava até escrevendo, nem acabei de escrever, mas eu tô ficando aqui em real time.

Uma maneira, tipo, que eu acho que é muito importante, uma técnica seria tipo, one-on-ones. Então, tipo, ter conversas individuais com cada um, acho que ajuda a fazer esses alinhamentos de uma maneira até personalizada, barra não... Não, como que eu posso dizer? A palavra fugiu. Não é agressiva a palavra que eu quero dizer, mas de uma maneira que não expõe as pessoas.

Às vezes, por esses motivos, porque a pessoa é tímida, seja lá o que for. Então, tipo, você fazer esses alinhamentos pode ajudar, e aí isso alavanca pessoas independente de que área, seja área técnica ou outra coisa. Então, tipo, aplicando como? Comunicação.

E o último item, que eu nem tinha chegado a ver e que eu tava pensando, seria... Não sei, acho que não. Acho que, tipo, entrar em algoritmos seria entrar em detalhes muito técnicos, assim, especificamente. Pensando no escopo disso aí, eu acho que vai muito na individualidade de cada um.

Então, não sei se aplicaria nesse contexto. Acho que entraria numa outra esfera, sabe?

Entrevistadora: Então, assim, o que você considerou é muito sobre a aplicação dentro do seu próprio grupo de trabalho. E não nas pessoas que vão ser impactadas.

Foi mais, assim, de dentro do grupo, entre os desenvolvedores.

Bruno: Sim, mas pensando por essa perspectiva também... Ficaria um pouco difícil pensar, tipo, fazendo um comparativo dessa visão que eu tive, específica de grupo, pra externalizar isso às vezes com pessoas que você vai estar interagindo. Pensando como se fosse clientes, seria essa perspectiva?

Entrevistadora: É, não tem certo também, tá? Assim, é pra entender como é um checklist pensando, fazendo um depara de um desenvolvimento de produto, né, pra um desenvolvimento de algoritmo.

Então, isso seria, fundamentalmente, pensando no impacto. Mas se a interpretação que você fez é pro grupo, né, seu grupo de trabalho, seus pares, tá valendo também.

Bruno: Não, a interpretação inicial que eu fiz foi em cima disso.

Mas, tipo, extrapolando pra esse outro universo, né, que eu fiz. Isso seria, tipo, seria da interação de quem a gente iria atingir, né? Seria a nível de cliente. Essa que é a intenção do resto, até dos outros casos, até pra contextualizar.

Entrevistadora: Tá.

Bruno: Não é? É isso?

Entrevistadora: Ah, sim. Não, na verdade, a princípio foi pensando em quem vai ser impactado. Tá.

Samaga: Eu que interpretei errado, então.

Entrevistadora: Não, não é errado. Mas, assim, de repente tá escrito errado. Isso. Mas, tá bom. Mas, obrigada. Renato, você quer falar?

Renato: Sim. É, eu pensei mais do lado de uma solução.

Mais do lado do dado, na verdade. Pensando que, muitas vezes, a gente assume que essas atividades são meio fixas, né? Assim, eu faço um cadastro pelo CPF e a pessoa não vai mudar de CPF. Mas a gente tem muita situação de coisa que a gente meio que considera fixa que pode mudar, né? Então, por exemplo, uma mudança de gênero.

Então, eu pensei um pouco nesse sentido, né? E aí vai um pouco do que eu entendi também já no último item lá de como que isso poderia ser aplicado. É justamente pensar no dado, nas informações que vão ser tratadas em algo que, por mais que, às vezes, intuitivamente a gente ache que é algo que nunca vai mudar e não tem que ser tratado de forma diferente, mas pode ser alterado. Então, nesse caso aqui, o nome da pessoa pode mudar por se ela tinha algum problema com o nome dela, por uma mudança de gênero mesmo.

O gênero, a forma que ela gosta de ser, que ela quer ser tratada, quer ser chamada. Então, eu pensei meio que nessa linha e aí como que isso entraria ali no desenvolvimento de algoritmos seria considerar que o dado que eu tenho no dia 1, quando eu fiz o cadastro, não necessariamente vai ser constante até no

ciclo de vida daquele usuário dentro do produto, né? Isso ao longo do tempo pode mudar. E aí, eu tô pensando numa recomendação.

A recomendação talvez tem que levar essas mudanças em conta também.

Entrevistadora: E aí você acha que, tipo, esse stop aqui, esse direcionamento pode te ajudar a, pelo menos, pensar? Porque, às vezes, pode ser uma coisa que você não está considerando inicialmente.

Renato: É, eu acho que é algo que talvez não venha naturalmente, assim, de já pensar, de explicar. E aí, tendo um direcionamento desse, ajudaria já a mapear e já partir do dia 1 com isso em mente, né? Que vai facilitar também esse desenvolvimento.

Entrevistadora: Legal, legal. Caio. Caio.

Caio: Eu trouxe para o segmento que eu trabalho. Então, assim, isso é seguro puro. Seguro, daqui a um ano, é diferente. Mudou a idade, mudou a idade do que está sendo segurado, a idade do condutor, mudou a idade do veículo. Então, tudo isso altera, né? A precificação do seguro, assim.

Quando a gente considera o tempo também, é muito importante. Por exemplo, a gente passou por um momento de COVID. Então, como que a gente coloca o que aconteceu na COVID dentro de uma estimação de modelo? Por exemplo, dentro do algoritmo para performar, para construir um cenário em que isso não seja impactado.

Então, basicamente, a solução que a gente encontra é controle. A gente controla isso para poder tentar tirar o sinal, tirar a influência dessa perspectiva, tanto de tempo quanto de mudança, quanto do que a pessoa era e do que ela vai ser, porque o risco também altera. Então, eu posso estar com um monte de carro no meio da minha vigência do seguro.

Então, tem uma mudança de perspectiva ali no meio e isso altera a precificação, isso altera a análise de risco, vai alterando tudo. Então, a ideia que a gente tem é sempre neutralizar. Pega esse sinal, difunde esse sinal para as outras variáveis, tenta entender como que as outras variáveis se comportam com esse novo sinal para a gente ter perspectiva dentro da modelagem, mas não ser dependente dela.

Até porque eu modelo para trás, mas eu precifico para frente. Então, não necessariamente é o que está acontecendo ou o que vai acontecer para frente. Então, eu trouxe mais para essa parte, para o campo que tem de experiência na parte de seguro, na parte de precificação de seguro.

Entrevistadora: Interessante isso. Muito bom. E é uma coisa que você sempre, no seu dia a dia, já é uma questão que você sempre pensou ou sempre foi? Já fez parte do desenvolvimento?

Caio: Sim, isso faz parte do desenvolvimento. É a primeira coisa que a

gente pensa. Quais são as variáveis que a gente vai controlar? O porquê que vai controlar? Uma variável também que não está aqui, mas é uma variável de flutuação também. É quando a pessoa muda de PJ para PF ou de PF para PJ.

Também o comportamento, a interação. E quando a gente olha também o tempo, o lifetime velho dessa pessoa, entra em bola todos os conceitos, porque não necessariamente ela pode ficar na empresa, ela pode alterar. Não é um produto.

Então, eu tenho um multicálculo de produtos que podem ser muito parecidos. Então, o preço decide muitas coisas. Essa perspectiva sempre está inclusa dentro da parte de precificação.

Entrevistadora: Legal. Afonso.

Afonso: Oi. Sobre o que eu entendi. Eu interpretei mais como você entender melhor a pessoa da sua solução. Eu entendi mais assim.

E está totalmente relacionado com as minhas atividades. Não totalmente relacionado com as minhas atividades, mas do projeto que eu participo, porque envolve dashboards. E quando envolve dashboards, você não conhece a sua pessoa, o seu público-alvo ali, quem está utilizando. Isso influencia 100

Então, para que uma solução nova, para eu fazer algo diferente, se aquilo funciona. Então, uma cabeça muito mais conservadora. Então, isso não é tão comum dos membros do meu time, mas foi algo que eu percebi.

Eu tenho alguns problemas de engajamento, e isso foi percebido nas reuniões. E foi algo também que, ao passar essa mensagem para o chefe da minha equipe, o gerente da minha equipe, a gente viu a necessidade de estar mais presente junto com o cliente, não ficando muito afastado de home office. Então, isso gerou até algumas viagens diferenciais para o cliente, para os usuários ali, para ter esse contato mais presencial, que é algo do perfil deles.

E isso não tende a mudar, até porque são pessoas já de 50 anos, algo assim. Então, não é de hoje que vai mudar o soft skill dele e dela. Então, surgiu essa necessidade de adaptar algumas ações nossas, até envolvendo um pouco mais de custo, mas, em troca, ter um engajamento melhor é algo muito benéfico para o projeto.

Então, é algo que a gente vem realizando no dia a dia e vem aplicando. E a gente percebe isso em reuniões. Então, a gente percebe que, por não estar presencial com a pessoa, não ter aquela resposta rápida, não ter aquela relação mais forte.

E com essas viagens que a gente fez, a gente viu que melhorou muito a resposta. A gente conhecia a face da pessoa pessoalmente, almoçou junto. Então, isso não muda muita coisa.

Então, eu entendi mais nesse sentido de conhecer melhor a persona do seu projeto. Ou com relação à adversidade, muito mais com a resiliência daquela pessoa. Porque o público do nosso projeto é muito mais homens, mais velhos, com uma cabeça mais conservadora.

Então, não tem muita adversidade nesse meio que eu atuo para poder influenciar o projeto em si, o desenvolvimento.

Entrevistadora: Entendi. Você pensou mais nas maneiras que as pessoas, o público já pode utilizar.

Que é aquilo que você tem e nas relações ali, diferentes maneiras, diferentes formas de se relacionar com o produto, especificamente.

Afonso: Exato. Se a pessoa é mais jovem, o dashboard pode ser feito de uma maneira diferente, talvez com mais informações.

Uma pessoa de mais idade, talvez menos coisas maiores para chamar mais atenção da pessoa, para não confundir. Então, isso tudo tem impacto no dia a dia.

Entrevistadora: Francisco.

Francisco: Vou ser breve aqui. Acho que... você tocou nos dois pontos que eu acho que também são um pouco parecidos. Eu também pensei no trabalho atualmente, pensei na questão de produtos financeiros, de crédito.

E sempre com aquela preocupação de você ter uma maneira de explicar as alterações. Por exemplo, explicar suas predições. Como é que eu explico para o Banco Central Europeu, Banco Central do Brasil, que um crédito não foi aprovado para essa pessoa? Se ela fizer uma reclamação, eu tenho que fazer essa explicação.

Se houver uma mudança de dado, alguma mudança de atributo nessa pessoa. Eu estou, mesmo que indiretamente, isolando as variáveis que não devem ser aplicadas para determinadas áreas. Então, saber isso, muitas vezes depende muito mais do que da nossa arquitetura.

Depende mais de um cuidado e de tal. Realmente, pensando nisso. Geralmente, não é o foco principal de uma empresa ambiciosa, pode não ser o foco principal.

Então, é de se pensar. É o impacto em algo que você está fazendo ali, que você está otimizando. Você tem que analisar como isso vai se comportar, se isso vai acontecer, isso também vai continuar estático com o tempo.

Nunca é uma coisa... Geralmente, não é uma coisa autônoma. Não é um monolito que você vai soltar lá para sempre. Então, eu penso mais nesse sentido mesmo de acompanhar o funcionamento do algoritmo, mais até do que... mais ainda do que a concepção.

Entrevistadora: Entendi. Como que, de maneira orgânica, ele vai ser absorvido e vai impactar? Exatamente. É mais por esse lado que você pensa?

Francisco: Isso. E se eu tenho os mecanismos para garantir a legislação, para avaliar essas variáveis sensíveis que eu colocaria, eu não sei, aí eu depender da aplicação, uma variável vai ser sensível em outra área, não vai ser... E se eu estou considerando isso aí, eu tenho como provar, se eu criei um mecanismo que eu posso ter isso. Ou seja, se eu estou observando, eu tenho que ter isso. Se eu não estou observando, muitas vezes eu não estou.

Então, se eu estou estruturando para que eu possa voltar no tempo e explicar isso aqui que aconteceu. Por exemplo, a gente falou ali de mudança de gênero. Isso deve impactar um crédito financeiro na pessoa? Em qual país? Como é a legislação para isso? Vai ter legislação para tudo isso, né? Então, é poder acompanhar tudo isso e monitorar para você levar em consideração até para questões legais, né?

Entrevistadora: Legal, gente, muito bom. Aquelas colocações. Eu vou para o segundo ponto aqui, eu vou fazer a mesma coisa, tá? Vou ler e aí se vocês puderem preencher. Eu acho que vou diminuir um pouquinho o tempo porque ficou três minutos.

Vou botar dois, que aí o que der para preencher está bom. Se não der, a gente discute. Identifique viés relacionados à capacidade, gênero, raça e ou acessos e preveja possíveis interações incorretas que possam acontecer.

Dá uma musiquinha. Não foi da outra vez.

(...)

E aí, pessoal? Pode falar, Du.

Afonso: Eu puxei primeiro, Larissa, porque eu achei muito parecido com o anterior. Eu tive esse entendimento, tá? Então, eu vou ser bem breve aqui. Tem até um caso interessante, interessante de um modo negativo.

De trabalhar com alguns clientes que você percebe um comportamento diferente quando está lidando com uma mulher ou com um homem. E através dessas reuniões, a gente percebe essa falta de confiança, principalmente quando lida com mulher. E quando entra um homem ali, parece que ele sente mais confiança no que a pessoa está dizendo.

Então, eu já senti isso na pele. Mas, no meu caso, um ponto positivo. Porque eu estava com esse homem.

E isso influencia totalmente na hora de direcionar as reuniões. E é um pouco difícil de lidar com isso. De como você vai chamar a atenção do cliente sobre isso.

Você meio que acaba recuando, talvez pegando essa profissional e colocando em outro lugar, não sei. Então, é um ponto delicado de falar, mas eu

vejo muito relação com os argumentos que eu utilizei no card anterior, no 01. Tá.

Entrevistadora: Mas nesse... Só para a gente fazer uma discussão breve sobre isso, aqui está fCaio da diversidade das perspectivas, né? E aqui está fCaio de... Já fez um recorte dessa diversidade, né? Você pensou de maneira já direta no 01? Você já trouxe todas essas questões?

Afonso: Exato.

Entrevistadora: Já te remeteu a essas questões específicas ali, tá? Legal. Depois foi o Caio. Pode falar, Caio.

Caio: Eu fui direto, assim que eu li, na perspectiva... Também que trouxe em volta a perspectiva anterior do meu trabalho, mas a minha cabeça foi direta na questão do que a mulher dirige pior do que o homem, né? E que isso não se dá nos dados, e é uma perspectiva muito... Assim, se a pessoa vem com esse preconceito estabelecido, quando ela olha os dados, ela quebra totalmente, ela acha que os dados estão errados, mas quando você olha o histórico, sempre foi assim. E existiu uma pesquisa, um tempo atrás, considerável, não vou conseguir afirmar quando, mas foi da CNSEC, que é um dos homens reguladores, de entender o porquê que as mulheres traziam menos sinistro do que os homens.

E existia uma hipótese estabelecida na sociedade de que a direção feminina era mais agressiva, ou dirigia pior do que a masculina. E constatarem algumas coisas de atenção, de que o jeito que fazia as manobras, o movimento era menos brusco, a condução era mais consciente, a atenção era redobrada em algumas situações. Então, eu acho que a gente traz isso... Na verdade, o algoritmo vai inverter essa lógica.

E, como estatístico, a gente sempre trabalha com hipótese. A gente sempre levanta esse tipo de hipótese para entender se é realmente isso que está acontecendo. Toda vez que chega algum caso pré-estabelecido para a gente, a gente entende que o mais recente agora não é nesse sentido de perspectiva humana, mas trouxeram falta de capacidade de carro elétrico, por exemplo.

Então, é uma hipótese que... É um viés que trazem... É o corretor, mas carro elétrico é ruim, bateu, acabou, destruiu tudo. Então, traz esse viés para a gente. A gente sempre senta, cria as hipóteses, entende o que está acontecendo, vai para os dados, e aí, sim, toma uma decisão para entender para onde vai, qual é o direcionamento, qual é a decisão.

E essas decisões são muito gerais. Depende muito do que a gente quer, mas, normalmente, define uma metodologia, entende qual é o caso, define uma metodologia, aplica e tem uma decisão.

Entrevistadora: Interessante, Caio. Davi, pode falar.

Davi: É, bom, eu já não consegui, na minha interpretação, fazer tanto

esse paralelo com o primeiro cartão. Nesse, assim, sobre o meu entendimento, que a gente deve sempre levar em consideração que os dados que a gente está usando, o nosso processo de modelagem, pode estar gerando algum viés, que, de alguma forma, pode ser prejudicial para uma pessoa, um grupo de pessoas, de forma indevida.

Trazendo um pouco, agora, aqui para o meu lado, não muito tempo atrás, saíram alguns estudos, realizados pelo MIT, sobre algoritmos de reconhecimento facial. Eles eram altamente racistas, principalmente para mulheres. Eles pegaram um estudo com o Senado dos Estados Unidos e fizeram uma busca de reconhecimento facial com todos os senadores americanos e uma base de procurados do FBI, e em torno de 60

Então, isso é uma coisa que é muito importante, pelo menos na nossa área, que a gente tenta levar bastante em consideração. Então, se é realizado no dia a dia, com certeza. Como você trocaria com as mesmas buscas, é sempre buscando entender quais são os possíveis vieses, que aquele algoritmo, aquele processo, aquele produto que você quer desenvolver, ele pode conter, e fazer uma ampla análise dos seus dados, do resultado dos seus modelos, para validar essas hipóteses, de que você está tendo um viés ou não.

E tentar, caso você tenha, empregar alguma técnica que seja apropriada para essa correção. Uma busca de mais dados, mais diversos, de classes mais específicas, gêneros mais específicos, ou diminuir, caso não seja possível, uma redução do que você está, que é majoritário, coisa do gênero. Basicamente foi isso que eu entendi desse cartão.

Entrevistadora: Fazendo um resumo do que o Caio falou, e você também, Davi, são questões, assim, que estão já latentes. Ultimamente, realmente, tem muitos estudos comprovando, fCaio, de várias questões do dia a dia, de diversidade, desses vieses que vão afetar essas questões mais sensíveis. E o Caio falou, então, de questões assim, tipo, quase que os dados, eles refletem mais a realidade, e isso ajuda a mitigar os vieses.

Assim, pensando pelo lado, como você deu o exemplo ali da direção, né? Existe esse viés inconsciente geral de que a mulher dirige mal, mas os dados estão refletindo outra questão, então, de certa forma, você até usa isso para mitigar, né? E o que você está fCaio também, Davi, é uma questão que é quase que assim, não tem como não pensar, né? Esses, por exemplo, essas questões de racismo, né? Até semana passada teve um caso, que eu acho que vocês devem ter visto, né? De um cara que foi identificado no estádio, no estádio de futebol, e foi quase preso, né? E foi super constrangido, enfim, coagido e tudo mais, e não era porque um cara negro, enfim, né? Então, são questões, e talvez, assim, aí eu queria até ouvir um pouco de vocês, se vocês

conhecem, sei lá, de uns anos para cá, se existem metodologias e ferramentas e barreiras que têm facilitado a não reprodução desses vieses, que está mais popular, né? No dia a dia, assim.

Davi: Acho que se fala mais sobre isso, né? Quando você traz visibilidade, você, de certa forma, você consegue apontar alguns dedos que incomodam, né? Então, por exemplo, tiveram muitas Big Techs que saíram desse ramo de reconhecimento facial, né? IBM saiu, tiveram outras que saíram também, e as que ficaram, criaram áreas de, né, IA for Good, que são especializadas em buscar problemas nesses produtos, né? Eu acho que do lado de técnicas, pelo menos não que eu esteja muito ciente, mas, com certeza, se está olhando mais para isso.

Entrevistadora: Tá. Só citar um exemplo rápido, né, que eu, na minha pesquisa aqui de doutorado, eu tranquei em 2022, e quando eu voltei, eu fiz um mapeamento, uma revisão sistemática, né? Só em 22 tinham, com um recorte super ali em cima da questão de vieses de gênero, tinham sido publicados mais de 180 artigos, e a grande maioria comprovando, fazendo teste, né, assim, matemático mesmo, para comprovar esses vieses, né? Então, eu imagino que para vocês que desenvolvem, isso deve estar cada vez mais latente, de fato, né?

Afonso: Um outro ponto, Larissa, cumprimentando também o que o Davi disse, eu acho que é muito mais a resposta da sociedade em relação àquilo que foi entregue, acho que isso impacta mais. A ferramenta em si, tem até uma que eu trabalhei uns cinco anos atrás, na mesma empresa que o Davi, né, que analisava o sentimento das pessoas envolvidas nas notícias, tá? Então, se acontece algum problema, a notícia é sobre o lançamento de uma tecnologia X, e com base nessa ferramenta você consegue entender o sentimento que o mercado teve, que as pessoas tiveram, nos principais meios de comunicação, em relação àquele tema, né? E isso vai impactar, se você bem, ou dar uma resposta, né, dar um parecer, ou deixar de dar continuidade ao projeto, ou adaptar ao projeto, né? Então, esse meio é uma ferramenta também de você conseguir entender o quanto aquilo foi impactante, positivamente ou negativamente, na sociedade para tomar uma ação no seu projeto, na sua solução.

Entrevistadora: Sim, faz sentido. A própria sociedade vai regulando, né?

Afonso: Exato.

Entrevistadora: Vou passar para o terceiro, tudo bem? Então, o terceiro aqui, crie uma diversidade de caminhos para participação em sua experiência. A verdade é que eu achei aquela chata, tá? Chata também.

(...)

Tá mais rapidinho agora. O que vocês acharam, pessoal? Vocês acharam, entenderam, acharam aplicável? Davi, você levantou a mão ou é das antigas?

Acho que ficou, né?

Davi: Perdão, ficou, sim.

Entrevistadora: Se quiser falar.

Davi: É, posso falar. O que eu entendi aqui é que, caso a gente esteja, né, desenvolvendo um experimento que necessita de algum tipo de participação de algum público, a gente deve se preocupar em prover diversos meios ou formas para que diferentes pessoas, com diferentes necessidades, possam participar desses experimentos. Acho que também, trazendo um pouco para o meu lado, está um pouco fora do que, do que eu trabalho, assim, no meu escopo atual, mas pensando como poderia, né, fazer isso, é entender quais são as dificuldades e barreiras que essas pessoas que eu gostaria que participassem do meu experimento poderiam, né, encontrar e tentaria arrumar soluções para elas.

Basicamente seria isso.

Entrevistadora: Renato.

Renato: É, eu pensei mais ou menos nessa linha do Davi, né, mas pensando no, no que as pessoas estão procurando na experiência, né, então, nos interesses que a pessoa pode ter, nas restrições do que ela tem ali.

Um exemplo que eu tenho em casa, né, minha esposa tem aracnofobia e ela sofre muito com, assim, estar usando o YouTube e, de repente, aparece um thumb lá de uma aranha gigante. E essas ferramentas, às vezes, elas têm alguma coisinha para você bloquear, uma palavra, uma coisa assim, mas elas são muito falhas, né, passa muita coisa, é muito complicado. Então, eu pensaria algo nessa linha, de dar para o usuário as possibilidades, tanto ele colocar o que ele tem interesse, o que que ele está procurando, mas também o que ele não quer, o que ele quer evitar, o que pode ser algum problema para ele, um gatilho, qualquer coisa do tipo.

Eu acho que isso é algo possível de ser feito, mas uma coisa que eu anotei aqui é que deixa tudo um pouco mais, o desenvolvimento mais complexo, pode aumentar custos, né, e aí pensando numa empresa, às vezes, vai ter um bCaiço ali, né, do quanto que a empresa quer investir nisso em relação ao que ela vai ganhar, digamos assim. Eu arrastei aqui, fiz uma bagunça. E, é, acho que é isso aí, como buscaria seria justamente nessa linha, né, de abrir para o usuário também poder dar o input dele ali, de como que ele quer experienciar a experiência que está sendo disponibilizada para ele e personalizando de acordo com os gostos e as necessidades dele.

Entrevistadora: Afonso?

Afonso: Eu fui para um lado um pouco mais técnico da coisa, tá, é, porque dependendo de como do seu produto, de como as pessoas acessam o

produto, se define muito bem o seu público, assim. Então, muita gente tem um produto que, usando alguns termos técnicos aqui, só pode ser acessado via chamada HTTP da sua API, né, mas se você desenvolver um dashboard amigável de pessoas que conseguem obter a mesma resposta, só que de uma forma muito mais amigável, né, você começa a atingir outros tipos de pessoas, né, e isso impacta totalmente o seu desenvolvimento. Então, em projetos que eu participo, como são pessoas, homens mais velhos, conservadores, no sentido de tecnologia em si, de usar tecnologia, a gente sempre tem o foco de deixar o mais prático possível, de clicar duas vezes num arquivo aqui, que vai tudo acontecer, né, o que talvez não seria necessário se fosse um outro tipo de público, né, por exemplo, ter um dashboard mais qualificado ali, mais esteticamente, se for um outro tipo de pessoa, né, então isso influencia no dia a dia, né, na forma daquelas pessoas acessarem, né.

Então, como a pessoa acessa hoje o seu serviço, a sua solução, isso impacta totalmente o público que você vai atingir, né.

Entrevistadora: Francisco?

Francisco: Bem rapidinho, basicamente eu também, eu pensei direto também na coisa de como é diferente lançar um produto, uma feature e acompanhar os dados, como é diferente você também desenvolver junto com pessoas diferentes, né, então, e você como pessoa muito diferente, nada semelhante, né, de preferência, é uma experiência muito diferente, você vai ter coisas que você usa todo dia, você nunca vai se perguntar se aquilo ali é interessante, tem um amigo que é deficiente visual e, pô, convivendo com ele, eu percebi tanta coisa que é, que foi modernizado, que foi excluindo, né, que foi excluindo, desde, sei lá, pra gente, pra mim é ótimo chegar no meu, tem três elevadores no meu prédio, apertar um botão e chamar qualquer um dos três, mas para ele isso aí é um problemão, porque ele não sabe onde vai chegar o elevador dele, por exemplo, né, desde a entrada, entrar em casa com os leitores capacitivos, né, de posicionamento variável, muitas vezes, né, então, assim, realmente você não tendo contato com o diferente, a chance de você fazer um produto também inadequado, né, mesmo para uma pessoa que tem uma questão de capacidade, né, mas estar próximo é uma ajuda, é muito importante para a gente diminuir o erro, né.

Entrevistadora: Abrir as perspectivas também, né. Muito bom. Bruno, quer falar alguma coisa?

Bruno: Eu ia comentar, eu tô meio perdido, na real, mas, tipo, pensando pelo ponto de diversidade de caminhos e participação e experiência, eu não sei, até coloquei aqui pensando em alternativas para atender diferentes públicos, né, beleza, mas aí depende do foco do que seria o objetivo ali do projeto,

do aplicativo, seja lá o que for fazer, de, será que você quer realmente ter diversidade ou quer focar em um nicho, sabe, e dependendo, falhar rápido seria um caminho bom, tipo, e aí você alteraria dependendo, quer aí a diversidade de caminhos aí para possíveis soluções dependendo sobre demanda, sabe, de feedback e não dependendo, ficar pensando antes em possibilidades para poder implementar, sabe, tipo.

Entrevistadora: Mas por que que você está perdido? Você fez uma boa colocação aí, parece, você não está entendendo.

Bruno: Porque lá em cima do objetivo estava, tipo, o primeiro, o primeira etapa do que eu fiz, eu até pensei, relacionado até ao time, né, nem pensei, eu me confundi, fui para o lado sem ser de produto, né, do cliente, mas até dependendo, em relação até ao objetivo lá, tipo, de soluções de impacto social, sabe, daí depende, impacto social em qual nicho, sabe, pensando nesse caso de diversidade, tipo assim, não poderia tacar para todos ao mesmo tempo, dependendo, sabe.

Entrevistadora: Ah, mas esse, em relação a esse um que você falou, a interpretação que você fez, ontem, a maioria pensou igual a você, eu acho que está dúvida, sabe, o jeito que está escrito.

Bruno: É, por isso que daí, tipo, depois desse primeiro, daí, dando o caminhado das etapas que a gente tem, até com a fala de cada um, sei lá, tipo, o meu parece que saiu fora da curva e aí bagunçou.

Entrevistadora: Não, mas é importante. Bom, vamos passar para o próximo, é, só fazer um comentário, o Davi e o Bruno tinham me falado que podiam ficar só até as seis, ficam à vontade, tá, se puderem ficar, maravilhoso, se não puderem, eu entendo super, podem ficar à vontade.

Bruno: Eu ia falar com vocês, mas eu, eu, eu, eu tenho no máximo, eu tenho seis e dez, na moral mesmo.

Entrevistadora: Tá, vamos, vamos então pular para o quatro aqui, fiquem tranquilos aí, tá? É, quatro, assuma que o projeto terá incertezas, necessidade de ter ações e adaptações e trabalhe com a utilização de habilidades complementares.

(...)

Vocês acharam desse? É, fazer um comentário, achei interessante, Bruno, que você falou, de repente, assumir que tem que errar rápido para ir ajustando, né? Acho que veio um pouco, né, sobre isso nesse, nesse item.

Bruno: Sim, é, tipo, principalmente de voltar da software, eu acho que isso é natural, sabe? Tipo, falhar o quanto antes é a melhor coisa possível que a gente pode fazer, porque, tipo, precisa do feedback, porque o software rodando, mas não sendo utilizado, não adianta de nada, porque é só dinheiro

gasto.

Então, tipo, ter o feedback, o quanto antes de quem usa realmente, para a gente poder trabalhar os pontos fracos ou, sei lá, as melhorias que tem que fazer, é o melhor caminho. E incertezas é relativo, tem que trabalhar em cima das certezas, das incertezas você vai, com o feedback, provavelmente vai trabalhar em cima. Então, a parte de adaptação, essas coisas em cima, acho que é uma adaptação natural mesmo, tipo, fazer o quanto antes, botar e ver quais são os resultados.

Entrevistadora: Uhum, bom. Caio?

Caio: As nossas interações aqui, às vezes, são muito complicadas, porque nós somos uma equipe estranha de técnica e vida, por exemplo, com corretor. Então, para a gente explicar o porquê que a gente está gravando em 15

Então, toda vez que a gente tem que lidar, assim, com áreas pares, que a gente tem a interação, mas que sejam de, não nível, mas assim, skill diferente, sabe? Extremamente complementares, nada de coisas que não têm interseção, a gente sempre faz uma prova de conceito, né, antes, qualquer tipo de projeto. Então, assim, é uma prova de conceito inicial, pensa todo mundo junto, o que é que vocês querem, depois refina isso, depois que refina isso, passa para um processo, assim, isso aqui a gente consegue fazer, dentro daquilo que foi pedido, aí eu digo de estruturação de modelagem mesmo, de software e tal, e aí a gente tem uma entrega final para o cliente, mas pensa muito por entender o que eles querem e eles entendam o que a gente está entregando, porque muitas das vezes acaba que essas interações, adaptações confundem mais do que necessariamente agrega ali em alguma coisa. Então, a gente faz esses filtros para poder trazer uma perspectiva mais assertiva, mas não enviesada, ou não estabelecida, sabe?

Entrevistadora: E considerando também o entendimento, assim, dos próprios clientes, ou as pessoas, dos usuários, né, como um conhecimento importante para agregar com o de vocês, né, que é mais técnico.

Caio: Exato, porque diferente de muitos produtos do seguro, ele tem uma interseção, né? Eu não falo direto com o meu cliente, eu tenho uma pessoa ali no meio, que é o corretor. Então, o que eu passo para o corretor, o corretor passa de acordo com o entendimento dele com o cliente. Então, eu preciso deixar isso claro, para que isso possa transpassar com uma naturalidade melhor, que o cliente entenda o porquê que está acontecendo aqui.

Por exemplo, tem muita blindagem. O cliente blinda o carro, acha que vai ficar mais barato. Não, fica mais caro, porque o carro fica mais pesado.

Quando bate, tem diversos problemas, fica mais caro para a oficina, então o risco fica maior. Como é que você explica que o seu... você está fazendo com

que você fique mais seguro, mas o seu risco fica maior? Então, assim, essas explicações são muito...

Entrevistadora: Entendi.

Caio: São muito complicadas para o corretor.

Entrevistadora: Contraditória, né? Alguém quer colocar alguma coisa? Eu acho que esse item é uma coisa mais até já muito do dia-a-dia, né, de vocês. Talvez essa questão já de pensamento ágil, metodologia ágil, errar, acertar, chamar pessoas, né, para participar também essa questão da POC, né? Enfim, acho que um pouco mais dominado, né? Pode falar, Francisco.

Francisco: Eu pensei um pouco, me veio na hora a questão um pouco mais focado na questão do algoritmo, né? E também de produto, né? Então, exatamente isso também é... Quando a coisa é incerta, abre margem para a gente fazer... a gente não saber o que a gente vai fazer, né? E aí é muito legal você poder, junto com o seu cliente, você chegar a entender, né? Por exemplo, hoje a gente atende lojistas e tal.

O que que o lojista quer? Ele talvez possa não saber o que ele precisa, o que ele quer. Ele pode chegar, ah, eu quero uma máquina de pagamento, mas será que a gente junto não vai entender que, na verdade, ele quer expandir o negócio dele? Ou será que nesse país a gente quer expandir o negócio? Nesse outro ele quer produtividade, nesse outro aqui ele só quer proteger a marca dele, né? E aí abre margem para você utilizar os dados e os experimentos para, junto com eles, né? Que você não poderia fazer isso, você não entende disso, né? Vocês chegarem a somar essa colaboração aí, realmente, para entender o que seria legal de fazer.

Entrevistadora: Bom, legal isso. Bom, vai lá, Bruno, obrigada, viu? Muito obrigada.

Bruno: Não sei se eu ajudei muito, mas qualquer coisa que você me chame, eu ajudo, tá? E aí, revisa aí, sei lá, ver o que tem que fazer.

Entrevistadora: Pode deixar. Te chamo para escrever, se quiser.

Bruno: Não, só se for código mesmo, mas tudo bem. Valeu, gente, valeu.

Entrevistadora: Tchau. Eu vou passar, então, aqui para o cinco, tá? Desenvolva uma alfabetização básica em acessibilidade e se aprofunde nos critérios de acessibilidade específicos relevantes para suas soluções.

(...)

E aí, gente, isso aqui é questão de acessibilidade, até você comentou um caso, né, Francisco? É um mundo, né? Um mundo de questões, assim, né? Para todas as... não sei nem... às vezes a gente fala errado, né? Mas, tipo, as questões que cada um tem, né? As sensibilidades ali de cada um, visuais e, enfim, tem várias... várias especificações.

Vocês já tiveram oportunidade de trabalhar em algum projeto assim? Tiveram que aprofundar? E vem isso no dia-a-dia de vocês? Pode falar, Afonso.

Afonso: Eu não tenho que contribuir, mas tenho que contribuir, que habilidade física, sensorial, não vejo sendo considerada, sabe? Essa é a minha visão hoje. É muito jogar de lado isso. E é importante também que Francisco comentou, né, dessa habilidade complementar, você tá em uma equipe multidisciplinar, pessoas de formações diferentes, pessoas com capacidades visuais diferentes, te faz enxergar, né? Porque, às vezes, tá muito enraizado em você que você nem pensa nisso, entendeu? Somente quando surgiu um caso desse, ah, eu tenho um amigo aqui que tem uma capacidade visual limitada, ele se depara com o que eu tô fazendo, o que eu tô entregando, e ele percebe isso, porque ele sofre com isso no dia-a-dia, né? Aí sim que eu vou começar a pensar nisso. Ah, realmente, cara, não tinha nada a ver com pensar nisso. Aí você começa a rever as coisas que você fez e nunca para pensar nisso, né? Então, essa é a minha contribuição e não contribuição ao mesmo tempo.

Entrevistadora: Tá muito distante ainda, né, do nosso dia-a-dia, assim. Um exemplo, recentemente entrou uma colega que tem deficiência visual num grupo de WhatsApp. Então, qualquer coisa que a gente... a gente, poxa, a gente é muito de imagem, né? A gente compartilha imagem o dia inteiro, e aí a gente tem que escrever, sabe? Deve ter ferramenta hoje que lê a imagem para ela e tal, mas é interessante esse exercício da gente se colocar no lugar do outro mesmo e ver que tem uma maneira totalmente diferente de interpretar aquilo que está acontecendo, né? E aí é isso, contato com um diferente que faz a gente pensar.

Afonso: Isso, eu só vejo quando você quer aquele público-alvo bem específico, aí você para pensar naquilo, mas na diversidade não vejo muito isso. Então, se eu estou desenvolvendo o celular para pessoas mais idosas, vou ter preocupação em botar letra grande, menos coisa na tela para... questão de cognitivo dos idosos, naturalmente é menor, né, do que os jovens, mas quando é seu público-alvo, quando não é e você para a pensar nisso, acho que isso não acontece, pelo menos comigo, né? Na minha vivência.

Caio: Só para compartilhar uma experiência, e assim, mais da imprensa do que necessariamente o mínimo, a Liberty... A Liberty, ela foi vendida, né? Agora ela virou a HDI, então a gente não sabe muito bem como vai ficar, mas a Liberty em si, ela tinha um processo de... a gente chamava de grupos de acesso.

Então, ela tinha um grupo que cuidava de pessoas maiores de 50 anos, um grupo que cuidava das pessoas com problemas de qualquer tipo de deficiência e tudo isso era compartilhado com todos. A gente tinha reuniões, acho que

se não me engano eram trimestrais, e essas experiências eram compartilhadas. Então, por exemplo, teve uma coisa simples para a gente que é... elevador, por exemplo.

Reformaram o prédio, colocaram um elevador novo, o elevador não tinha sonorização. Quando a pessoa deficiente conseguia apertar o número, porque tinha o Braille, mas ela não sabia se ela estava para andar certo. Então, se entrava outra pessoa com ela, e apertava outra, por exemplo, ela apertou o 8, a pessoa apertou o 4, ela saía no 4. Então, assim, essas coisas pequenas que é fácil de resolver de muitas maneiras, assim, a gente acabava tendo contato e ficava mais fácil da gente conseguir achar uma solução para isso.

Entrevistadora: Legal mesmo, muito legal. Isso enriquece muito, né? E, assim, a tecnologia vai lá na frente, e às vezes a gente não pensa na coisa muito básica, né, de inclusão. Impressionante. É... Então, vamos para o último. Alguém quer colocar mais algum ponto? Acho que foi, né, esse... é um ponto pacificado que a gente precisa melhorar, né, como um todo. E, então, o último aqui.

Adote uma definição mais flexível de projeto, abra os processos e convide pessoas com habilidades relevantes a contribuir com seu projeto, mesmo que não sejam programadores ou programadoras. (...)

E aí, pessoal? Pode falar, Afonso.

Afonso: Caraca, isso aí tem... algo que me chocou.

Eu fui me aventurar com o PowerBI, né, fazer algum dashboard. Eu sempre tive a parte do back-end muito forte e front-end nada, né? E aí, fui fazer um dashboard, eu olhei vários sites ali para me basear em um dashboard bonito de PowerBI e tudo mais. E, para mim, estava, pô, bem legível, bem legal.

Aí, quando eu fui mostrar para o meu irmão que a parte que lê de comunicação também fez design e tudo mais, ele olhou para aquilo e, tipo, não entendeu nada do que eu estava fazendo. E, para mim, está muito claro. Os títulos dos gráficos, está muito claro.

E ele falou assim, fCaiodo muito engraçado. Como assim, você não está entendendo? Estou fazendo isso aqui, ó. E o usuário, bem ou mal, também, geralmente, não é na área de TI, de programação, o usuário, né? Então, isso me ajudou bastante a... tá, então. Aí, fui lá e conversei com ele o que ele entendeu ali e adaptei, né? Mas foi engraçado essa diferença de percepção do que eu estava fazendo de outra pessoa que não era da minha formação, né? Até uso, algum dia, alguns termos mais técnicos, assim, que, para mim, está tão natural que para os outros também vão entender, né? Mas, na verdade, não entendem.

Entrevistadora: Mas você falou assim para o seu irmão.

Afonso: Foi muito problemática a situação.

Entrevistadora: Você chegou para o seu irmão e falou assim, deixa de ser burro! Como assim você não está entendendo? Foi assim?

Afonso: É, mas só parecia que eu não conheço. Como você não está entendendo aqui, ó, o gráfico assim? Você não entende o que está acontecendo isso?

Entrevistadora: É, é legal, né? Ver as outras perspectivas mesmo, assim, e ajuda bastante, assim.

Afonso: É, e sobre como eu aplicaria, não vejo outra alternativa senão envolver um pouquinho essas pessoas, nem que seja pontualmente, né? Como esse foi o meu caso, que eu mostrei o dashboard para ele, para ele enxergar que eu me considero um pouco mente aberta, assim, só que tem alguns casos que realmente não passam na minha cabeça. Com esse caso do elevador, por exemplo, para mim é tão natural que não passa nessa cabeça, né? Somente quando acontece, assim, eu ouço falar que é, realmente faz muito sentido.

Então, o que o Caio falou também fez muito sentido, é algo que eu nunca parei de pensar, né? E se eu desenvolvesse uma solução sem ter alguém com essa perspectiva, talvez cometeria o mesmo erro, entendeu?

Entrevistadora: E mais do que, assim, a gente sabe que tem essa coisa já incutida no desenvolvimento de ir testando, ir mostrando, né, durante o caminho antes de estar pronto, mas mais que isso, quando eu falo aqui na questão de abrir o projeto, abrir os processos, né, porque às vezes lá no comecinho o jeito que você está estruturando o seu processo pode, se você discutir isso com alguém, né, se discutir com outra pessoa que tem outra visão, isso pode contribuir também, né? Não esperar ter algum resultado, né, de fato, especificamente nisso. É isso.

Renato: Eu até coloquei aqui, né, de incluir essas pessoas com outros backgrounds em sessões de brainstorming no começo do projeto para ajudar a entender os problemas e planejar, né, como que eles vão ser tratados. Porque como, como que eu buscaria, né, lidar com com essa demanda.

Entrevistadora: Muito bom. Eu queria que você, agora, não sei se alguém quer colocar alguma coisa sobre esse item, não quero cortar não.

Eu queria fazer uma última interação com vocês aqui, é uma votação, vocês podem votar em todos, mas a ideia é assim, votar nos que vocês acharam que são mais relevantes desses seis, eu vou abrir a votação aqui, e se quiser colocar um, dois, três, ou não, os seis são relevantes, pode colocar, e eu queria depois ouvir um pouco também o que vocês acharam no geral em relação a um ao outro, né, se eles se repetiram, o que vocês acharam. Deixa eu liberar a votação aqui, aí, colocar a área, abriu aí para votar? Negocinho, botãozinho

de baixo aqui.

(...)

Todo mundo votou? Ou estão pensando ainda, né? Sem pressão.

Afonso: Ah, tem que colocar aí, I'm done, né? Ok.

Entrevistadora: Boa. E aí, o que vocês acharam? Já vai aparecer o resultado aqui? Acho que só eu que não votei, porque eu estou me abstendo aqui.

Vocês acham que se repetiram, a temática, né, em alguns cards se repetiram, dá para ser mais direto, e vocês acham que foi relevante, vale a pena trazer essa reflexão?

Afonso: Para mim, eu achei o 1 e 2 muito semelhantes. Talvez o 1 é um pouco mais abrangente, mas para mim ficou muito na mesma natureza, assim. Essa foi a minha visão.

Posso estar errado na interpretação, provável que seja isso, mas achei muito parecido ali, entendeu? Acho que outras pessoas aqui também, na hora de comentar, comentou isso a respeito. Eu perdi o resultado aqui, mas ele fica registrado, depois eu acho que vocês votaram mais no primeiro, né? Ou no segundo? Não vi, gente. Acho que foi 1 e 3, né? 3 votos.

Afonso: Vou tirar uma foto aqui.

Davi: O terceiro é o mais votado. Junto, empatado com o primeiro.

Entrevistadora: O terceiro está empatado com o primeiro, tá.

Davi: Todos os outros têm dois votos.

Entrevistadora: Você achou os dois primeiros? Beleza, foi o 3. O primeiro recebeu 3, o segundo 3, tá.

Depois empatou todo mundo. Pessoal...

Davi: O primeiro recebeu 3 e o terceiro recebeu 3. Tá. Acho que tudo dois.

Entrevistadora: E aí, assim, vocês acharam que alguma coisa trouxe para vocês refletirem diferente que vocês já estão acostumados? Acharam muito utópico? Vocês acharam? Querem dar um resumo final?

Davi: Eu achei o 3 bem interessante. É uma visão que, pelo menos da forma que eu interpretei, eu nunca tinha pensado antes. Nunca tinha parado para pensar nisso.

Os últimos 3, para mim, são um pouco fora da minha realidade, assim. Embora eu ache importante essa questão levantada no 6, no 5 e no 4, mas um pouco fora da minha realidade. E o 1 e o 2, acho que já estava dentro do que já está dentro do meu dia a dia.

Mas esse 3 realmente foi bem diferente.

Caio: Na minha visão também, o 3 é o mais diferente para mim. Eu até cheguei a comentar que os 5 são coisas que me fazem pensar, mas me fazem pensar no sentido de melhorar alguns processos.

Não de iniciar. O de iniciar, para mim, é o 3 e talvez seja o que eu tenha tido mais dificuldade na hora de desenvolver algum pensamento.

Entrevistadora: Legal. Interessante. Você também achou, Francisco?

Francisco: Sim, compartilho também. O 1 e o 2 já é um ponto na área de dados que já é latente.

Do 4 ao 6, mais focado em ações. Achei também que o 3 é o mais reflexivo.

Entrevistadora: Você vai falar, Afonso?

Afonso: Na minha visão, é muito relacionado ao engajamento do usuário em relação ao produto.

O 1 e o 2, para mim, é o que mais me impacta, é o que mais me preocupa. E também ouvi algumas pessoas aqui falarem da questão da acessibilidade. É algo que, no meu caso, eu muito descantei e foi importante ouvir, porque isso abre a mente.

Eu gostei bastante disso, foi bem benéfico.

Entrevistadora: Legal. Renato?

Renato: Eu acho que o 3 é, para mim, o principal e acho que ele acaba um pouco cobrindo os outros.

Muito do que foi discutido nos outros, o 3 ajuda, de certa forma, é um caminho para se resolver. Então, você tem um problema de diversidade do 1, por exemplo, e você ter uma solução que vai ter... aí pensando um pouco de como eu interpretei o card mesmo, mas ter uma solução que vai dar liberdade daquela pessoa usar aquele produto, ter uma experiência adequada de como ela quer ter, isso já vai resolver um pouco de dificuldades que ela tem nos outros. E aí, para a questão de acessibilidade, a mesma coisa, você teria que resolver a acessibilidade para permitir que uma pessoa que precise de acessibilidade seja atendida por esse card 3 também.

Acho que ele acaba cobrindo a maior parte dos outros cards, poderiam ficar um pouco embaixo dele.

Entrevistadora: Verdade. Ele dá uma resumida, vira um guarda-chuva.

Legal. Gente, é isso. Muito obrigada, eu sei que final do dia, todo mundo cansado aí, todo mundo super participou, muito obrigada mesmo.

Eu não falei no começo, mas os dados de vocês estão preservados, tá? Eu não vou usar o nome de vocês na pesquisa, nem imagem, fiquem tranquilos. E eu gostei muito da discussão, vocês trouxeram exemplos práticos muito bons. Com certeza eu vou colocar no meu texto, vai ajudar bastante.

E é isso. Obrigada, gente. Boa noite para vocês.

Afonso: Parabéns pelo trabalho, Larissa.

Entrevistadora: Obrigada.

Francisco: Muito legal.

Todos: Obrigada. Obrigada a vocês. Tchau, tchau.

Tchau, pessoal. Prazer.

C.2

Grupo de foco (mulheres) - Checklist para o desenvolvimento de algoritmos sensíveis

Entrevistadora: ...porque o automático dele é inglês? Foi. Não, peraí. É, foi.

Então, gente, vou compartilhar, te dei a autorização, Renata, ver se você conseguiu. Foi? Ótimo. A gente passou agora pela apresentação da dinâmica.

Todo mundo já mexeu no Miro? Pode falar, Rebeca?

Rebeca: Eu não.

Alice: Eu nunca também.

Entrevistadora: Não?

Leandra: Não, mas parece o Mural, né?

Entrevistadora: É, isso, parece bastante.

Você não levantou a mão, não, Rebeca? Confundi e tá, beleza. É, para dar zoom é aqui na rodinha do mouse, tá? Então, aqui assim, é, com o Tab, né? Tab faz a mãozinha, se arrasta para lá e para cá. Aqui do lado tem a barra de ferramentas e aí você coloca, você insere texto, insere elementos, notas, assim, tipo post-it para escrever, enfim, é bem intuitivo.

E aí a gente vai passar agora pelos itens, os seis itens que eu falei que eu acho que são legais a gente discutir aqui. E aí depois cada uma tem o seu quadrinho com seu nome aqui, tá? Então, vocês podem, eu vou explicar, vou passar pelos itens, aí vocês podem, cada uma ir para o seu quadrinho e a ideia é que, assim, a cada leitura de cada item tem algumas perguntas, tá? Assim, o que você entendeu, é, se você acha que é realizável no seu dia a dia, se faz sentido, né? Como você aplicaria, quais meios buscaria e se esse item te fez pensar algo que você não costuma pensar, geralmente, na hora do desenvolvimento. Essas perguntas a gente vai passar com mais profundidade quando a gente estiver discutindo aqui, só falando.

Mas, é, a ideia, eu vou dar três minutinhos para cada item, para vocês só colocarem, dar uma refletida antes da gente começar a discutir. Aí, se vocês quiserem escrever aqui alguma palavra para depois vocês lembrarem, né? Para jogar na discussão ou alguma questão, vocês queiram chamar atenção, algum ponto de atenção, enfim, podem colocar aqui. É, não se preocupem com o

tempo, não fiquem tensos aí, porque vai ser pouquinho para cada um, que o legal mesmo é a gente se esgotar na discussão, tá? Mas, aí, isso aqui é um guia para a gente.

Então, é, começando aqui, o primeiro item seria esse. Eu não vou explicar por enquanto, tá? Porque a ideia é vocês pegarem e maturarem aí o que que é. E, aí, eu vou ler para vocês. Considere a diversidade sob a perspectiva das interações humanas e considere a maneira como as pessoas mudam ao longo do tempo, tá? Então, esse é o primeiro item.

Aí, vou pedir para vocês irem lá para o quadrinho de vocês, achar ali o item 1, e vou botar aqui o cronômetro, três minutinhos para vocês dar uma maturada e ver se entenderam, se não entenderam. Se não entenderam, depois, quando a gente começar a discutir, vocês falam, gente, falou grego, tá bom?

Renata: Larissa, Larissa, tem algum tipo de permissão especial no board? Porque o meu não está aparecendo essa barrinha aqui do lado, não, para escrever.

Leandra: Para escrever, falar isso. Não estou conseguindo escrever.

Entrevistadora: É, qual barrinha? Aqui?

Renata: Não, essa do lado, essa barrinha de ferramentas.

Entrevistadora: Ah, que menina.

Leandra: Só está deixando a gente selecionar ou abrir um comentário.

Renata: Ou selecionar um comentário.

Alice: Ah, eu achei que era para comentar.

Entrevistadora: Eu coloquei como editoras. Eu vou... deixa eu mudar isso aqui.

Alice: Agora foi.

Entrevistadora: Ah, achou. Ah, que bom.

Renata: Era isso mesmo.

Entrevistadora: É, era isso. Então, vou soltar, tá? A primeira rodada. Qualquer coisa, vocês me chamem. Vou botar a musiquinha. Se estiver ruim a musiquinha, vocês me avisam que eu tiro aqui.

Renata: Tem quanto tempo para cada item, Larissa? Você separou?

Entrevistadora: Para toda a fileira do item 1, vão ser 3 minutos, tá? Vai ser muito rápido, mas aí é só a ideia inicial, comecei ali, para maturar mesmo o pensamento e depois a gente discute. Vocês estão escutando a música?

Alice: Ah, me atrapalhei para descer. Como que mexe mesmo para descer no Tab?

Entrevistadora: É, você aperta o Tab, aí vai fazer a mãozinha e você empurra com o mouse.

Estou aí aqui. Quase, né? Faltam 5 segundos. Vocês conseguiram dar uma refletida? Então, eu queria agora discutir, a gente vai ter mais ou menos uns 10 minutos, mas aí a ideia é a gente tentar estressar, estressar no sentido, né? Trabalhar bem os pontos que forem surgindo. Queria perguntar, começar perguntando para vocês, se você, se alguém não entendeu esse item e se não entendeu, se realmente assim ficou complexo e tal, por que que não entendeu? Todo mundo entendeu, conseguiu alcançar?

Alice: Eu acho que eu entendi, mas eu não tenho certeza se eu entendi o que você queria que eu tivesse entendido.

Rebeca: É sobre isso, eu acho que eu entendi, mas o que eu entendi é o que você gostaria que a gente tivesse entendido, Lari.

Alice: Eu queria ter entendido, não sei se entendi corretamente.

Entrevistadora: Minha ideia é entender, né? Entendi.

Então, mas o que vocês entenderam? Não tem problema, tá? Se estiver totalmente errado, a errada estou eu, que escrevi o negócio ruim. Então, não tem errado aqui. Alguém queria dizer, assim, o que que foi o primeiro entendimento, né? Dentro daquelas perguntas que a gente fez ali, o que que entendeu mesmo sobre o item? Alguém quer fazer um resuminho?

Leandra: Eu botei aqui sobre... Ah, a Renata botou a mão.

Renata: Já começou, pode começar. Eu fico só com a mãozinha aqui, depois eu falo.

Leandra: Mas eu acho legal usar a mão aqui.

É só, eu entendi a questão e levei mais para o lado pessoal, né? Até coloquei aqui, que além de ter diversidade no time, é importante considerar que as pessoas podem mudar seus pensamentos e atitudes. Então, eu entendi esse primeiro ponto, como é importante levar pontos pessoais das pessoas, né? Ver que elas são pessoas, não são só recursos e ir trabalhando para desenvolver algo.

Entrevistadora: Só uma pergunta Leandro, depois eu passo para a Renata.

Você então considerou que essa mudança das pessoas seriam as pessoas que vão receber, que vão ser impactadas pelo algoritmo, ou são as pessoas que estão junto com você desenvolvendo?

Leandra: É, junto comigo, dentro do time. Tá. Mas aí talvez, não levei para esse ponto das pessoas.

Entrevistadora: Legal, é um bom ponto de vista também. Renata, quer falar?

Renata: Tá, vamos lá. Eu vou falar assim dos dois, que é, dá uma apreensão de forma bem geral, que ninguém é estático, né? As pessoas mudam.

Então, talvez o comportamento que você pegou hoje de uma pessoa não vai ser o comportamento dela de amanhã, ou do ano que vem. Aí você pergunta, né, se você quer passar por todos os itens ou você quer o item por item, Larissa?

Entrevistadora: Pode ir, pode ir. Se você quiser, se fizer sentido, pode ir.

Renata: Pronto, porque aí cadê o pensamento? Aí tem uma pergunta que você fala se acha viável. Aí eu coloquei um depende, porque eu acho que isso depende muito. Depende, pensando em algoritmo, tanto de dado que você tem sobre o comportamento da pessoa.

Pensando em interação do dia a dia é quanto você interage com alguém, sabe? E no final, que você pergunta o quanto isso pode impactar, isso me levanta um questionamento. O quanto que é difícil você usar algoritmos para prever comportamento. É muito difícil, não é fácil.

Porque toda vez que você prevê o comportamento de alguém, você está olhando para o recorte das ações mais recentes dela. Isso que é feito em algoritmos como da Netflix, das redes sociais. Elas meio que olham o nosso comportamento mais recente e começam a jogar um monte de coisa para nós.

Só que comportamento é algo muito mais amplo. Não só o que eu fiz recentemente, mas o que é um comportamento de risco que aí você extrapola para os algoritmos de fraude. Assim, é difícil responder isso.

É difícil monitorar. Então, eu acho que olhando essa questão da diversidade, das mudanças dentro dos algoritmos, eu encaro como um desafio. E eu não tenho uma resposta para te dar.

Esse é o melhor jeito. Mas sei que é desafiador. E que talvez você precise de muito mais embasamento para conseguir trabalhar com o comportamento de pessoas que mudam o tempo todo.

Entrevistadora: Embasamento, você diz, assim, questões psíquicas ou comportamentais de profissionais que tenham embasamento no problema específico. É isso que você diz ou é uma coisa mais técnica?

Renata: Eu falaria uma coisa mais técnica. Você prevê o comportamento de alguém.

Eu falo embasamento, mas na verdade você tem que ter muito mais informação para dizer o que uma pessoa vai fazer. Ou o que é um comportamento de fraude. Ou o que é um comportamento de um possível comprador.

É difícil você dizer, porque tudo é um recorte. Mas aí entra. Será que eu tenho mais informações? Ajuda melhor nos meus algoritmos? Então, olha um pouco sobre essa perspectiva.

Entrevistadora: Bom, Rebeca, pode falar.

Rebeca: Muito interessante o que a Renata falou. Eu fui mais para o lado pessoal também, pensando ali como a Leandra que falou primeiro.

Entendendo que a diversidade gera um desenvolvimento pessoal em quem está trabalhando junto. Mas nesse sentido de desenvolvimento. Tem uma pesquisa que é muito interessante, Lari.

Ela fala que as pessoas tendem a ser a média das pessoas que mais convivem com ela. E se eu tenho uma diversidade convivendo próximo de mim eu vou tender a me desenvolver muito mais pessoalmente. Eu não pensei muito no algoritmo nesse plano.

Entrevistadora: Você também pensou nesse sentido da equipe, da sua equipe de trabalho.

Rebeca: Exato.

Entrevistadora: Em relação às pessoas que você vai trocar, vai discutir, debater.

Rebeca: Exato. E eu acho que a gente não tem muita diversidade no nosso ambiente. A gente é uma fatia da sociedade só.

Não reflete a sociedade. Então, eu acho que é isso.

Entrevistadora: E aí fica um pouco pasteurizado na hora de...

Rebeca: Obviamente que a discussão no nosso ambiente é rica. Mas não é uma diversidade que reflete a sociedade.

Stepanie: Galera, o meu virou uma mistura aí. Com o meio pessoal, meio profissional.

Então, vamos lá aqui. Sobre essa parte do ambiente ser diverso. Responde parecido com as meninas.

Principalmente de como vai mudando... Que nem o time que eu estou hoje, a maioria era homem. Começou comigo ali de mulher. Então, você vai vendo que no começo eles não queriam te ouvir.

Querendo ou não, acaba sendo machista. Tem uma certa trava, mas vai mudando. Você vai conquistando ali seu espaço.

Eles acabam até baixando a guarda. Você vê até com outras mulheres quando começam a envolver. O próximo ponto é de praticar no dia a dia.

Uma coisa que indo para o lado pessoal... Eu me vejo fazendo hoje em dia principalmente com as redes sociais. É acompanhar outras pessoas diferentes. É PCD, pessoas que têm outros tipos de sexualidade. É começar a entrar em contato. Você pode não ver ali fisicamente, ao vivo. Mas o acompanhar você faz pensar sobre outras coisas.

Por exemplo, você falou... Eu pensar no próximo me faz considerar ele. Então, por exemplo, esses dias eu estava na rua. Parei no farol e tinha um poste... Espera aí que eu vou abrir a porta.

Ele abriu. Tinha um poste onde uma cadeira de roda desce. Então, se eu não acompanhar as pessoas que falam sobre isso... Por exemplo, às vezes eu

nem me tocava sobre isso.

Então, isso cria desníveis. O que você está... O que você convive? O que vocês se expõem? Acaba modificando a forma que você vê as coisas. Foi por esse lado que eu fui na análise.

Entrevistadora: Pode falar, Alice.

Alice: Mais um complemento mesmo do que as meninas já falaram. Eu acabei tendo um enfoque mais pessoal.

Menos indo para a parte de algoritmo. O que eu tinha entendido a princípio... Era mais entender a diversidade como uma potencialidade para todo mundo. A gente sempre fala sobre diversidade.

Pelo menos isso é algo que eu estou acostumada. Mas quando a gente realmente para para olhar em volta... Nós estamos inseridos num ambiente diverso? Nós estamos frequentando coisas diferentes? Nós estamos, sei lá, lendo mulheres negras de outros países? Ou a gente está sempre consumindo e lidando com as mesmas coisas? E falando de diversidade de maneira mais superficial. Então, o que eu pensei de ser realizável, eu acho que sim.

Mas para isso a gente também tem que fazer uma escolha. Nós temos que ser assertivos nesse ponto. Porque a diversidade também tem que partir de uma auto-reflexão.

Do que a gente faz, do que a gente convive. Então, eu acho que é possível a gente lidar com potencialidades, com essas mudanças. Desde que a gente reflita sobre onde a gente quer inserir isso.

Se é só num algoritmo, se é na nossa vida. Mas eu acho que tudo isso passa por um caminho de reflexão, de percepção como indivíduos. Não acho que é algo que seja construído só por osmose.

Acredito que a gente tem que ter uma reflexão mais profunda. E realmente refletir onde a gente quer chegar. Então, se a gente talvez está querendo construir um algoritmo, não sei.

A gente tem que refletir no impacto disso e em como nós queremos que ele seja. Então, tem que pegar o barco, o remo e nós mesmos fazer esse processo. Que nem sempre é algo simples.

Acho que foi mais por esse lado.

Entrevistadora: Eu achei interessante, só até fazendo um resumo. Pelo que eu entendi, a maioria de vocês foi por esse caminho de entender. De olhar a diversidade nas relações entre a equipe de desenvolvimento. E como as próprias pessoas se modificam. E aí trouxeram questões a Regina, a Alice.

De você se modificar fazendo a abertura de um olhar mais amplo para as outras diferenças. Para as outras questões que estão na sociedade. E também esse primeiro... A colocação da Renata também achei muito legal. Que bate

até em uma questão da tecnologia em si. É realmente muito difícil você tirar padrão de comportamento. São tantas situações, tantas pessoas, dados.

Qualidade de dados, enfim. Mas eu achei interessante isso. Mas eu queria saber também se vocês conseguem ter esse tipo de reflexão. Pensando do ponto de vista de produto. Como seria isso, por exemplo. Ao fazer o desenvolvimento de um algoritmo. Você considera que ele vai impactar outras relações humanas. E pessoas que vão mudar ao longo do tempo. No sentido, qual seria mudar ao longo do tempo. Não só psiquicamente, mas por exemplo. Eu vou desenvolver um produto para uma criança. Essa criança vai ficar adulta, vai ficar idosa. Pode ser que fique. Então, considerar também essa passagem de tempo. Vocês acham que vale, de repente, até refinar esse indicador aqui. Esse item. Para fazer vocês direcionarem pensamento para o desenvolvimento em si. Não só... Porque em relação à equipe, ao pensamento. Eu acho que isso tudo... Tenho certeza que todos vocês vêm pensando. Vêm desenvolvendo. Mas claro, estou falando aqui de sete pessoas.

Mas em relação a pensar o impacto em si. Entendeu? Porque a equipe, a gente não tem muito... Por enquanto, a gente não tem muito o que fazer. Como a Regina falou. Ela entrou, só tinha ela de mulher. Ela não tem, necessariamente, tanto poder ativo de trazer a Rebeca. Vamos dizer, de ter trocado o gerente que agora é uma gerente mulher.

Enfim, são coisas meio que exógenas. Então, isso a gente não tem muito como alterar agora. Mas nesse pensamento de... Eu posso projetar algo que faça com que as pessoas se sintam mais confortáveis. E pessoas de diferentes aspectos. Diferentes interações humanas. Isso seria legal? Dá para interpretar dessa maneira? Fala, Ana.

Ana: Eu interpretei. Eu coloquei. Então, ali na minha resposta, eu coloquei... Eu coloquei a diversidade de gêneros. Neurodiversidade. Coloquei até tipo de corpos, raça. E também a de idade. De deficiência também. Porque acho que a gente tem que englobar também. E eu coloquei mais pensando no caso de um produto em si.

Porque eu pensei assim... Talvez... Então, vamos desenvolver um aplicativo. Então, aquele aplicativo tem que ter uma usabilidade ali que seja interessante para uma pessoa mais idosa usar. Que vai ter mais dificuldade ali com a tecnologia. Ou então se eu vou desenvolver um outro aplicativo. Que vá, sei lá, um aplicativo didático. As crianças que vão usar.

Então, assim, a gente tem que atender ali as diferentes idades. E o que elas pedem para a gente. Mas a gente também tem que saber interpretar muito o nosso público. Por exemplo, se tiver um público neurodiverso. O que a gente faz. Enfim, dá para pensar em uma infinidade de coisas aí.

Que vão... Além um pouco da interação humana ali até no caso. Fiquei até me perguntando. Será que isso cabe aqui? Porque vai um pouco além, né? Mas também está ali dentro, né? Então, assim, sei lá, uma pessoa que é daltônica.

Então, a gente tem que ter todo o cuidado na hora que a gente está fazendo alguma coisa ali. Para saber se aquela pessoa vai conseguir ler. Vai conseguir interpretar todas as informações de forma correta. É... Mas eu acho que... Que é mais nessa parte, assim. Das pessoas para quem a gente vai desenvolver alguma coisa. Então, sempre pensando no próximo.

Mas é que nem muita gente falou aqui. A gente tem que pensar nisso. Se rodeando disso no dia a dia, né? Então, acho que quando a gente só convive com os nossos iguais. A gente meio que para de pensar. Que nem um exemplo do poste no meio da rampa de acesso. Então, assim. Eu tenho um filho. Então, eu sei o que que é. Quais são as dificuldades ali de uma criança. Porque eu vejo pelos olhos dele muitas coisas no mundo.

Meus pais são idosos. Então, eu vejo pelos olhos deles muitas coisas no mundo. E eu acho que é dessa forma que a gente vai vendo. Eu tenho um cunhado que ele tem deficiência. Então, eu vejo muita coisa pelos olhos dele. E acho que isso vai contribuindo também, né? Para a gente entender essas coisas.

Entrevistadora: Com certeza absoluta. Muito bom, gente. Eu vou... Acho que a gente podia continuar falando disso. Mas a gente vai... Vou passar para o próximo. Para a gente conseguir cumprir aqui. Mas está ótimo, viu, gente? Todo mundo contribuindo bastante.

O segundo item é... Identifique vieses relacionados à capacidade, gênero, raça e ou acessos. E preveja possíveis interações incorretas que possam acontecer. Tá? Do ponto de vista, assim.

Vamos colocar assim. De uma maneira até que o algoritmo que vocês estão desenvolvendo é um produto mesmo, né? Ele vai ter um impacto. Nesse impacto, você consegue identificar, né? Vieses relacionadas à capacidade, gênero, raça. E consegue prever essas interações incorretas. Vou dar três minutos aqui para vocês. Então, gente.

O que vocês entenderam? Esse eu acabei até dando uma explicadinha, né? Antes de dar o cronômetro aqui. De iniciar. O que vocês entenderam sobre... Pode falar, Renata.

Renata: Tá. Isso aí é um assunto que eu gosto muito de comentar. Particularmente. E acho que principalmente... Quando a gente fala... Eu vou falar muito de inteligência artificial aqui, tá, gente? Porque a minha área. É... No processo de tomada de decisão. Porque hoje a gente fala o tempo todo de IA generativa.

A gente fala de modelos que já estão aí na nossa vida. Estão decidindo muitas coisas. E aí o questionamento é... Se a gente refletir o mundo real... Que esse é um problema, por sinal.

Porque o mundo real... Ele é enviesado. O mundo real... Ele é complexo. E por quê? Primeiro. Muita gente falou aqui de equipe. Isso de alguma forma já insere um viés. Porque se eu sou de uma equipe... De um grupo majoritário... Normalmente eu vou inserir os meus estereótipos... Dentro das coisas que eu construo. Ah, Renata. Isso eu não faço. Será? Lança esse questionamento.

Eu vou sempre analisar o que você está fazendo. Será que você não insere? Então já começa. Você tem o viés da pessoa que está construindo. Tem o viés dos dados. Os dados são enviesados, gente. Se você analisar, por exemplo, dados bancários... Tem muito mais de homem.

Mas muito mais. E tem todo um contexto histórico... Por que isso acontece. Aí tem o viés de região geográfica, por exemplo. Eu morei na região nordeste, gente. Existem muitos estereótipos sobre a região nordeste. As pessoas acham que você mora no meio da seca.

Entendeu? Que tem um monte de caca ao seu redor. Que é assim que as pessoas pensam que a nordeste é. Tem um monte de caca ao seu redor. E está todo mundo passando fome. Então acho que... Se é a região nordeste... Existem já estereótipos na sociedade. E você consegue inserir isso até com os próprios dados reais. Porque os dados reais... Você tem muito mais transação bancária na região sudeste. Pode ter certeza que tem. Então, se você usar o mundo real, vai estar inserido isso lá dentro. De alguma forma.

Então como é que você consegue mitigar isso? Isso falando de alguns vieses, já que estão os dados que a gente usa. Aí vem outro. Por que eu tenho que identificar para o modelo que foi uma mulher ou foi um homem que fez aquilo? Qual é a necessidade? Ah, mas isso vai melhorar a minha acurácia.

Tá, beleza. Mas você acha que o mundo está... Você acha que... Isso realmente reflete o mundo? Ah, porque são os dados. Mas como se os próprios dados do mundo real já estivessem enviesados? Então fica esse questionamento. É desafiador construir algo que não tem um viés lá dentro. E eu digo que é difícil eliminar completamente o viés. Você vai tentar diminuir.

Aí eu vou dar outro exemplo. Os modelos de identificação de rostos, de tem viés lá dentro? Com certeza. Modelos de análise de exame de imagens.

Tem viés lá dentro? Tem. Porque as próprias amostras que as pessoas têm, que os pesquisadores têm, que estão construindo, tem viés. Então como é que você vai mitigar isso? O que é realmente necessário para um modelo? É só melhorar a acurácia? Fica esse questionamento aí para o pessoal.

É só o número que interessa? Eu discordo. Inclusive, quando você vai

para um modelo de produção, tudo muda. Outro ponto que eu acho que... Eles falam de viés de gênero. Eu tenho viés de idade também. As pessoas têm os viés dela. Não tem jeito. Tem um outro item que eu coloquei. Deixa eu pegar aqui o meu quadro. Eu não coloquei aqui, mas eu vou dizer porque eu lembrei.

Que é fazer tudo o que o cliente pede. Eu estava uma vez conversando com uma menina. Eu dei uma palestra um dia em um evento. E aí a menina assistiu e depois ela veio conversar comigo no privado. Ela não quis falar para todo mundo. E ela falou, Ah, Renata, então... Eu queria te fazer uma pergunta, mas eu não queria falar na frente de todo mundo.

É o tal XPTO no projeto que eu trabalhei e eu acredito que tem um viés de gênero. Eu falei, por quê? Eu primeiro perguntei o que é que está sendo usado nesse modelo. Ah, tem alguma coisa do gênero lá.

Identificação, né? Eu falei, você consegue descobrir isso analisando os resultados desse modelo. E aí você faz as relações. Quem preveu que está com X ou Y e qual o gênero dessa pessoa. Você vai descobrir que tem. E eu disse para ela, provavelmente tem. Depois você só avalida isso que eu estou falando para você. Mas faz uma analisezinha aí que você vai encontrar. Então... Essa preocupação também... Vou fazer tudo o que me pedem ou vou indicar a gente? Não é assim. Eu, se passar por mim, eu falo, olha, veja bem, eu discordo se você está querendo colocar aqui.

Você vai ser infiel. Posso até fazer por obrigação, mas não porque eu concordo. Eu vou informar claramente.

Tem sim, tá? E é bom chamar atenção a isso porque para algumas pessoas só refletir a realidade está bom. Mas a realidade não está boa. Então a gente pensar como pode equilibrar isso. Para não só... Porque os algoritmos podem piorar o que já está ruim. Se ele não for bem pensado. É isso, gente.

Entrevistadora: Mariana.

Mariana: O que eu queria dizer um pouco dos dois temas que a gente está falando é que é muito complexo, realmente. É até difícil para mim conseguir ter o entendimento correto ali do que... Onde você queria que a gente chegasse.

Em que linha seria a linha mais adequada de pensamento. Mas focando no segundo, que agora as coisas ficaram um pouco mais claras para mim. O que eu queria contribuir é muito do que a gente acabou de ouvir da Renata, que é o seguinte.

O quanto as informações pessoais são importantes para algoritmos, dependendo do seu objetivo final. Se a gente está falando de seguro de automóvel, se a gente está falando de chances de acidente, se a gente está falando de recomendação. O quanto informações pessoais são relevantes para a construção desse algoritmo.

Se for para melhorar o acerto dele, como ela falou, para melhorar a acurácia, beleza. Mas faz sentido considerar esse tipo de dado. O que eu interpretei aqui desse cartão, desse card, foi que, de fato, dependendo do produto a ser desenvolvido, questões inclusivas precisam ser pensadas.

Vamos falar, por exemplo, de pessoas com deficiência auditiva. Pessoa com deficiência auditiva precisa muitas vezes de legenda ou ela precisa de um intérprete ali para poder estar auxiliando em alguma determinada informação. Muitos sites já pensam nisso.

O YouTube tem legenda automática, cursos online estão ali tentando incluir essas pessoas também. Mas, ao mesmo tempo, é muito comum da gente ver que até você ter uma pessoa deficiente auditiva no teu curso, você não pensou em desenvolver uma legenda automática, por exemplo. Até você ter um cadeirante como um consumidor do seu produto final, você não pensou que era importante ter a rampa.

Então, é muito complexo. Eu estou aprendendo com vocês, porque muitas coisas, infelizmente, a gente só enxerga o que está dentro da nossa bolha. E, até mesmo, focando aqui na questão do algoritmo, eu, Mariana, por diversos momentos dos projetos que eu passei e das construções que eu passei, eu não enxerguei até o momento informações de gênero, por exemplo, como informações cruciais para determinada resposta.

A não ser, por exemplo, o seguro de automóvel. Existe, comprovadamente, por dados que pessoas do sexo feminino se envolvem menos em acidentes dependendo da idade, do que os homens. Então, é muito complexo para mim.

Mas o que eu quero dizer é o seguinte, eu até falei isso na entrevista que eu fiz com a Lari. Esse tipo de dados, ao meu ver, ele vai ser relevante quando a gente conseguir explicar. Explicar que faz sentido ele estar ali, entendeu? Como a gente explica no seguro de automóvel pessoas mais jovens tendem a ter mais acidentes, pessoas solteiras também, pessoas de sexo masculino também. No caso, a questão do sexo é diferente do gênero. Pessoas, comprovado em dados, pessoas do gênero masculino e solteiras têm a probabilidade maior de se envolverem em acidentes do que pessoas do sexo masculino casadas. Porém, no caso da mulher é o contrário.

Mulheres casadas, historicamente, se envolvem mais em acidentes do que mulheres solteiras. Tem coisa que a gente não explica. Mas o dado explicava em outra. Era um dado generalizado? Não. Era um dado de cliente daquela bolha que é aquela empresa, daquela seguradora. Outra seguradora pode ser que tenha um resultado totalmente diferente.

Porque ela atua em regiões diferentes, em classes sociais diferentes. Um

seguro, dependendo da empresa, de uma seguradora que vem com banco ou não, ele é maior do que o outro, mas ele tem mais benefícios. Então, assim, a gente sempre vai ter vies, que é o que ela falou. A gente sempre vai ter, porque a gente vai estar sempre olhando bolhas. A gente nunca vai estar olhando tudo. Dificilmente a gente vai olhar o todo.

Então, falei, falei, mas quis dizer que a questão de gênero, raça, ela é bastante sensível, pensando em algoritmos. Tem que fazer muito sentido você estar usando aquele dado ali, para que você não seja injusto com alguém ou com algum nicho devido a essa informação.

Entrevistadora: Achei interessante o que você falou. Para inserir, você tem que conseguir explicar. Só esse exercício já faz você pensar, expandir um pouco ali para justificar. Legal. Alice?

Alice: Então, é mais um complemento, assim, do que tanto a Géssica e a Camila já falaram, mas uma coisa que eu estava refletindo, antes de eu ir para essa área de TI, eu fiquei um tempo na parte de editorial, de livros. Sempre é o que eu gostei muito. E na leitura, a gente tem o que se chama leitura sensível.

Então, por exemplo, um determinado autor faz um livro e aí ele contrata pessoas de um determinado grupo, da minoria, que talvez seja abordado nesse livro, para fazer uma leitura sensível. Não é uma leitura sensível para que isso seja aprovado em um grupo. É para que, se esse autor tem uma responsabilidade, ele esteja lidando com essa responsabilidade de forma correta.

É um produto que se está enviando para o mundo, a partir do momento que você escreve, ele não possa ser mais seu. Então, eu acho que, talvez, um exercício similar na construção de algoritmos deva ser feito. Não no sentido sensível de lidar com dados sensíveis, mas de ter uma ótica sensível ao que estamos lidando, ao produto que está sendo construído.

Então, aqui, sobre ser realizável, vendo numa ótica de produto, confesso que não tem familiaridade. Então, eu acabo tendo meu viés aqui mais para a parte de humanas. Eu não acho que isso deva ser uma questão de ser viável ou não.

Eu acho que isso deveria ser realmente uma prioridade. Desde a concepção do projeto, todas as etapas, até o final, até o levantamento do que foi o projeto, eu acho que isso deveria estar em todas as etapas, não apenas nos dados. Os dados é até um pouco óbvio que a gente tem esse cuidado a mais para não replicar padrões, para que a gente não contribua ainda mais para o que já é enviesado.

Então, eu acredito que deva ser algo assim que também tenha uma coisa gerencial, que não fique apenas na estrutura dos dados, mas que seja algo intrínseco do projeto como um todo. Então, eu tenho um pouco mais dessa

visão mesmo. E sobre ser realizável, eu realmente acho que deveria ser uma prioridade, porque se a gente for parar para pensar, talvez a inclusão ou o olhar da diversidade só está ali quando é conveniente financeiramente, ou então quando tem algum sentido.

E não é isso que a gente quer. A gente quer realmente uma mudança ou dar um stop em algumas coisas, a não ser que seja o objetivo do produto. Eu quero fazer uma coisa nichada apenas para mulheres casadas.

Então, aí tudo bem. Mas fora isso, eu acredito que a gente tenha sim que levar isso na construção, desde uma coisa bem embrionária até o final para que seja uma prioridade, que nem é quando temos uma leitura sensível. As pessoas fazem o estudo e depois contratam alguém.

Então, é realmente a forma como isso vai ser levado para frente.

Entrevistadora: Criando etapas, dá para criar etapas de verificação também.

Regina: Acho que no final das contas, o que eu vou falar acaba sendo um resumo. A gente tem que pensar no que eu quero responder, o que eu quero entregar, para quem eu quero entregar, quem realmente importa eu me basear no quê, quais informações e como eu quero entregar. Até onde eu vou me desdobrar para atender o público. Qual é a minha necessidade? E pensar principalmente no básico, que nem deram o exemplo de ser daltônico, talvez aumentar uma fonte de uma letra, por exemplo.

São, às vezes, detalhes que fazem diferença. Um exemplo que eu queria também trazer, numa vaga de emprego, o pessoal vai selecionar os currículos. Às vezes, uma faixa etária de uma mulher, eles vão pensar assim, dos 28 ali até uns 30 e poucos ela vai querer ter filho. Eles podem acabar não querendo contratar uma mulher nessa faixa etária, já pensando por esse lado. Pelo Estado civil também. Então, só um exemplo de situação que pode acontecer.

Entrevistadora: Aí é isso. Muito bom. Pode falar, Rebeca.

Rebeca: Complementando a área, eu acho que quando a gente não tem esse convívio da diversidade, quando a gente está desenvolvendo um certo assunto específico, então, por exemplo, como a gente estava falando de PCD, eu acho que é fazer uma pesquisa antes, é buscar pessoas que convivam com aquilo para poder considerar dentro do algoritmo. Então, estudar antes de, né, se eu não sei nada sobre aquilo, eu tenho que tentar estudar e não só já sair desenvolvendo, acho que estudar um pouco sobre o tema.

Entrevistadora: necessário, né? Muito bom, gente. Já fico pensando aqui nas ideias, mas eu acho que, isso que você falou lá no começo, Mariana, é o jeito que está escrito, eu preciso amarrar melhor para ficar mais claro, né, o 1, com certeza que o item 1 e o 2 também, mas eu até pensei em sair e perguntar

aqui para vocês, vocês acham que tinha que ter um guia com todas as questões de, né, que a gente falou, de gênero, rasping, pode falar, Mariana.

Rebeca: O bom é que, só uma bobeira, o bom é que a Lari fez testar primeiro com a gente, né, ao invés de testar com os meninos, né, não sei por quê. Mas, Lari, eu acho interessante se deixar solto para ver como que vai indo, onde caminha.

De repente, fazer uma segunda rodada com as perguntas mais amarradas.

Mariana: É, eu acho legal o guia, porque dependendo do dia que você teve, também, sobre coisas, assim, pode demorar mais tempo do que quando você já está pré ali, alinhada do que vai ser a parada. Mas deixar livre tem a sua, tem o seu valor, né, tem o ponto positivo.

Mas eu acho que um resuminho, assim, pelo menos de, olha, suponha que você é uma pessoa desenvolvedora de produtos, etc, e meio que uma introdução para que a gente possa seguir a mesma... Ah, não sei também, porque pensar em outras linhas é legal, né? Uhum. É porque eu fiquei bastante perdida, assim, no início. Então, não sei se as outras se sentiram perdidas também, mas...

Renata: Larissa, você pode colocar o seu tema de pesquisa, assim, não direto, vai fazer, ah, inteligência artificial, desenvolvimento de software, coloca, assim, o tema genérico, que aí a pessoa vai fazer a junção da pergunta do tema com... Porque também, se você explicar demais, você vai inserir um viés, tá? Porque os homens vão ficar com vergonha, e vão falar, não, agora eu vou pensar nesse tema.

Não, você tem que deixar eles sem saber.

Entrevistadora: Por isso que eu não botei... Por isso que eu não botei eles juntos. O que é especial? Oi, desculpa.

Renata: Assim, tema genérico, desenvolvimento de software, sei lá, gerenciamento de equipe, qual é o viés, pra onde é que você quer ir? Mas não coloque o que é pra fazer, porque se você colocar, algumas pessoas vão ter vergonha, e vão falar que se importam, mas não se importam.

Entrevistadora: Não, com certeza. É por isso que eu não coloquei junto, porque eu achei que eles iam ficar muito direcionados.

Renata: É, você já percebeu que tem diferença já?

Entrevistadora: Já, com certeza. Vamos para o três, então. Crie uma diversidade de caminhos para a sua participação, para a participação das pessoas que vão receber ali o impacto do algoritmo em sua experiência.

Assim, uma diversidade de caminhos para participar da mesma experiência. Vou dar o tempo de vocês aqui, três minutos.

(...)

Entrevistadora: E aí, pessoal?

Alice: Essa foi mais difícil, na minha opinião.

Regina: Posso começar? Posso. A minha resposta foi no sentido de ter importância de incluir quem vai consumir durante o desenvolvimento, a gente incluir quem vai consumir, quem tem interesse também, pra gente principalmente focar nas dores da entrega, né? Entender melhor qual que é o problema, ter bastante feedback, né? Procurar não só um feedback histórico, mas ir conforme for dando continuidade o projeto, né? Acompanhando o que o pessoal está sentindo do produto. E é nesse caminho de trazer pessoas que importam e entender e ver o que está trazendo de resultado.

Entrevistadora: Só depois eu fiquei pensando aqui, talvez eu não tenha feito também uma introdução muito... Às vezes a gente fica tão no tema, né? Que precisa dar uma saída, assim. Mas é legal pensar também que cada um desses itens, né? É quase que já é um guia, uma espécie de guia, mas é um stop e quase um direcionamento, assim. Pense na diversidade das experiências, como você falou bem agora, né, Regina? Tipo, quase que se colocar no lugar de outras pessoas ali para entender as diferentes maneiras de usar um produto, de ser impactado, enfim, tá? Então é por aí mesmo.

E aí a gente está discutindo aqui muitas questões que surgem, né? Principalmente porque vocês têm experiência em desenvolvimento, mas a ideia é entender se realmente faz sentido também, tá? Assim, tipo, você se imagina, né, num desenvolvimento, tendo que pensar ou já passou por isso, né, de pensar a experiência, né, para conseguir ajustar bem ali o jeito que vai se mudar o algoritmo, o produto em si.

Regina: É, seria essa questão de entender as dores do processo, né? Ver o que realmente de informação ali importa no final das contas. A gente consegue ver ali quais são os dados que têm mais peso, quais informações a gente tem mais acesso, mas não quer dizer que só a teoria tem a teoria e tem a prática, né? A prática acaba trazendo mais qualidade, né, no final das contas.

Entrevistadora: Ah, bem legal o que você falou. Pode falar, Renata.

Renata: É, eu entendi esse aqui, nesse item sobre a interação do meu usuário com o meu produto. Posso estar errada, tá, gente? Mas depois você filtra aí o que você acha mais relevante. Mas qual o ponto que eu entendi esse item? Como o meu usuário interage com o meu produto? O que o meu produto traz para alguém, né? E aí você perguntou sobre ser realizável, falei que depende, eu adoro essa palavra. E como eu aplicaria, tá? Eu acho, primeiro, um ponto importante é que, às vezes, eu acho que as pessoas lançam produtos no mercado, lançam produtos para as pessoas sem olhar se o produto realmente está sendo relevante para elas.

Porque, às vezes, a única métrica que é utilizada é a quantidade de ven-

das, a quantidade de usuários. Mas tem outras métricas que são importantes, né? Quanto que o usuário usa e depois desiste? É o tal do churn. E se estão usando adequadamente? Se está sendo fácil de usar? Se não estão usando porque está difícil? E aí você engloba todas as diversidades de uso. Só que eu ia cair um pouco para isso, para fazer pesquisa.

Não as participantes da minha pesquisa. Tentar pegar o maior número de diversidade possível nesse uso, tá? Isso pensando no usuário ativo do meu produto. Só que eu falei da criação de uma métrica relevante porque eu acho que métrica é muito além de acurácia, precisão ou mesmo do número de vendas.

Eu acho que dá para criar uma métrica customizável e possa avaliar de verdade os resultados para aquilo. Quão, é... quão diverso, assim, quanto que ele atinge, que público ele atinge, o quanto, o quanto que ele desiste. Dá para criar coisas mais customizadas.

Claro, depende do problema. E aí o que esse item me fez pensar é que você lança algo e não avalia o que aquilo está trazendo para a sociedade. Você analisa realmente, falando de algoritmos preditivos, você analisa os resultados que ele está propondo ou você só faz aquele pipelinezinho bonitinho? E vai por métricas de sucesso e não avalia ele mais a fundo.

Que tipo de resultado ele traz? Isso falando em algoritmos preditivos, mas quando você fala do uso de sistemas, você avalia quem é que utiliza o teu sistema, como ele utiliza, ele está sabendo fazer e às vezes eu acho que as empresas não fazem. Então pensar um pouco isso, quanto que eu busco quem usa o meu trabalho, o feedback dela, quanto aquilo está sendo útil para ela. Então, me deu essa impressão. Pode ser a filtro, se estiver errada, eu senti.

Entrevistadora: Está ótimo. Quem quer colocar mais alguma coisa sobre esse item?

Rebeca: Eu posso colocar, Lari.

Entrevistadora: Pode, claro.

Rebeca: Uma coisa interessante que eu estava pensando aqui agora, pensando na diversidade de caminho, é porque às vezes também quem está nos gerenciando vai influenciar também no nosso resultado. A gente estava num projeto que o objetivo era atingir a meta de resíduos lá descartáveis, agora eu não me lembro exatamente dos resíduos que a gente estava fazendo. Isso, ter o zero, esse daí. E quando a gente começou, o nosso direcionamento é, não importa o histórico, a gente quer olhar a meta. E aí, eu sinto aquela coisinha de olhar o histórico.

E aí a gente na reunião, conversando com o cliente, veio falando, a gente tem que olhar o histórico porque como vai começar a ser uma meta, a pessoa pode pegar a área e descartar uma coisa que ainda não deveria ser descartada

para poder ter lá a quantidade de toneladas e atingir a meta. Então, eu acho que quem está em cima da gente, também gerenciando, também vai influenciar no caminho que a gente vai seguir.

Entrevistadora: Faz sentido, sim. Tá. Eu vou para o próximo, até que eu não quero passar tanto do horário de vocês, tá gente?

Assuma que o projeto terá incertezas, necessidade de iterações e adaptações, e trabalhe com a utilização de habilidades complementares.

(...) E aí, pessoal? É, o que vocês acharam desse? Falou um pouco aqui da ideia de metodologia ágil também, né, e de você ir compartilhando, um pouco de multidisciplinaridade também nessa questão de habilidades complementares. Vai lá, Renata.

Renata: Tá. Eu vou falar um pouco do que é assim, da forma que eu entendi, tá? A forma que as equipes lidam com a gestão da incerteza e estar fora do mundo ideal. Aí eu falo de desenvolvimento de software de uma forma geral. Eu acho que as equipes não sabem lidar bem com a incerteza. Então, tudo é planejado para o ambiente perfeito.

Então, temos prazos adequados, temos o tamanho da equipe, temos todas as pessoas com as habilidades necessárias, temos tudo perfeito, lindo e maravilhoso. No mundo real, o prazo tá apertado, não foi planejado pensando em alguma incerteza que pode acontecer, tá faltando recurso, se é projeto de dado ou não tem dado, tá tudo completamente caótico. Se é com equipe, assume-se que todas as pessoas têm todas as habilidades, mas na verdade não foi pensado nisso, nem todo mundo tem todas as habilidades necessárias.

Então, eu brinco que eu tive um professor na faculdade, eu gostava muito dele, que ele falava assim, era um professor de estrutura de dados. Ele dizia, gente, tudo tem o melhor caso e o pior caso. Pensem pelo pior caso. Nossa, Renata, você é muito pessimista, não? Você pensar em uma gestão de tempo, onde você tem um tempo hábil, caso aconteça alguma coisa, não tem um recurso, uma pessoa se demitiu, o ambiente caiu, tudo subiu de preço. Coisas que podem acontecer e que dá pra você, não é uma previsão, mas dá pra você planejar que isso pode ocorrer. E você lida melhor com a gestão dos expectativas do cliente.

Porque é aquela coisa, o cliente, quando você é entrega antes do prazo, ele fica muito feliz. Quando você é entrega fora do prazo, ele vai ficar muito chateado. Então, saber manejar isso, e às vezes eu sinto que não. Em vários projetos que eu já estive, o caos impera. O caos e o excesso de otimismo, que eu acho que é problemático, principalmente no desenvolvimento de software. Então, eu acho que saber lidar com isso, e acho que isso passa logo pelo planejamento.

De você pensar que não seja muito otimista, seja um pouco mais pessimista, mas planeje pra isso, pra saber lidar com essas questões. Essa é a minha visão do tema.

Entrevistadora: Até pra não suprimir as etapas importantes que a gente vem discutindo, inclusive. Pode falar, Ana.

Ana: A minha interpretação já foi um pouco mais pro lado do soft skills ali, né? Habilidades complementares. Ah, beleza. Então, tenho uma incerteza aqui. Então, vou ter que ficar um pouco ocioso agora. Então, tem que lidar com a paciência, tem que lidar com o gerenciamento ali de humores, de uma equipe, ou qualquer coisa do gênero.

E dizem assim, a gente ouve muito como mulheres que, as mulheres elas são mais emotivas, e não sei o que lá. Então, tipo, mas um homem contrariado, realmente, causa um estrago muito maior do que uma mulher. E uma vez, eu já ouvi isso de gerentes mulheres de que nós, mulheres, nós somos melhores no soft skills do que os homens, por isso que geralmente as mulheres sobem, elas sobem pras áreas gerenciais.

Eu não sei se foi o viés que gerou isso ou se isso que gera o viés das mulheres na gerência, se é porque as mulheres sempre foram gerentes de lares, enfim, né, na história da sociedade. Mas eu vejo mais nessa parte, assim, entendi mais nisso, nessa parte de então, ah, vão ter adaptações, então a gente vai ter que lidar com essas adaptações, tem que lidar com parar e pensar se a gente consegue. Então, tem que ter aquele ponto ali de parada e de reflexão, né? Então, tem que estar aberto também a não ser tão rígido em relação a qual caminho que a gente vai seguir.

Então, se a gente tiver que pensar em outras soluções e abrir mão da nossa solução e seguir com a solução de outra pessoa, por exemplo, né, essa briga de ânimos aí, não sei se eu dei uma viajada aí, mas foi isso que eu pensei.

Entrevistadora: Bom, achei legal trazer esse ponto, porque isso também afeta totalmente, né, assim, a troca que você tá tendo ali, como você é direcionado, é realmente importante. É um ponto que eu não tinha pensado, mas eu acho que é super válido, inclusive.

Até porque isso, né, se você também não se sente confortável para colocar a sua opinião, trazer alguma questão, algum questionamento, provavelmente aquilo ali vai ficar prejudicado, né, e vai, pode enviesar, pode, enfim, trazer um resultado ruim, né?

Mariana: Eu queria rapidinho só falar uma linha de pensamento que eu tive, que foi bastante diferente. Eu até escrevi aqui, Lari, talvez você depois vai entender, mas eu já olhei esse cartão pelo sentido de, será que o produto desenvolvido é, não, será não, o produto desenvolvido, ele vai precisar de

adaptações, porque nem sempre a pessoa técnica que desenvolveu o produto entende do negócio. Então, sei lá, por exemplo, se eu tô fazendo um produto para a indústria e eu não tenho vivência na área da indústria, como técnica que desenvolve modelos de previsão, modelos de probabilidade, pode ser que o resultado daquele modelo não esteja representando a realidade da indústria.

Então, ter uma pessoa, quando eu vi de habilidades complementares, eu pensei nas pessoas do negócio, que não tem habilidade técnica nenhuma, mas que pode contribuir com. Olha, não faz sentido essa informação gerada, não faz sentido o modelo achar, sei lá, que pessoas de turnos diferentes vão ter chances diferentes de se envolver no acidente, sei lá. Eu já olhei mais por esse lado, né, de muitas das vezes nós nos metemos no negócio para entregar produtos de inteligência artificial, mas aquele negócio precisa estar bem apresentado, né, ele precisa estar sendo representado pela realidade.

Então, eu olhei de uma forma diferente.

Entrevistadora: Essa inserção multidisciplinar, né, para... É um complemento.

Mariana: Nós, sem as pessoas de negócio, não conseguimos, muitas vezes, entregar algo que realmente faça sentido e seja grandioso para o consumidor final, porque a gente não está naquele mundo ali.

Principalmente consultoria, né, que a gente vê projetos de diferentes brands e tal.

Entrevistadora: Interessante. Várias interpretações diferentes aqui, mas todas com bastante impacto. Faz muito sentido.

Regina: É que nem, por exemplo, você está acostumado, né, que nem tem o caso agora que o pessoal está trabalhando com texto. Ah, você já é acostumado a usar tal biblioteca, mas não quer dizer que para outra situação, né, para outro cliente, faça sentido você usar aquela mesma forma de tratativa, né? Então, quem está avaliando também não pode estar enviesado ali, né, tem que estar aberto a ter outros contatos, né, ver outras possibilidades.

Entrevistadora: Um exemplo disso. Eu fiquei achando que vocês iam fazer um tipo de resposta antes de perguntas, de ter as entrevistas, aí vocês responderam outras coisas. É isso.

É... Vamos passar para o próximo ou alguém quer colocar ainda? Vamos, né? Tá. A gente pode falar, tá? Não estou querendo tolir vocês.

Alice: Eu estou indo para o próximo, Larissa, que já já vou ter que sair, mas aí eu vou terminar de responder tudo.

Entrevistadora: Já passou, eu até peço desculpas aí, vou acelerar agora no finalzinho.

Alice: Nada que isso, a gente imagina.

Entrevistadora: É... O cinco. Desenvolva uma alfabetização básica em acessibilidade e se aprofunde nos critérios de acessibilidade específicos e relevantes para as suas soluções.

(...)

Rebeca: Lari, eu tenho que sair.

Entrevistadora: Pode ir, vai lá, Rebeca. Obrigada.

Rebeca: Beijo, pessoal. Beijo, meninas. Tchau, gente. Depois, se você quiser, eu termino de preencher, Lari.

Entrevistadora: Tá. Se você puder, uma boa. Obrigada. Tá bom? Beijo.

Rebeca: Tá bom. Beijo.

Entrevistadora: Eu vou, eu acho que eu vou passar para o próximo e a gente discute depois no final os dois, tá? Quem puder ficar.

Pode ser. Boa. Aí o último aqui é adote uma definição mais flexível do projeto, abra os processos e convide pessoas com habilidades relevantes a contribuir com o seu projeto, mesmo que não sejam programadores.

Deixar aqui o tempinho também para preencher.

(...) Tchau, Ana.

Gente, eu queria até aproveitar, assim, quem tiver alguma questão específica para trazer, para a gente discutir, mas, assim, dando já um apanhado de geral, queria pegar um pouco da opinião de vocês em relação, assim, vocês acham que foi repetitivo? Se teve alguma coisa que foi repetitiva, dá para condensar um pouco mais esses itens? E alguma crítica pela escrita, né, que com certeza dá para melhorar, mas se vocês têm alguma coisa mais específica para dizer, tá? E eu queria fazer uma votação rápida também, aqui, dos quais itens que vocês acham mais relevante considerar e tentariam aplicar, né? Assim, talvez vocês levem aí para o dia a dia de vocês.

Aí eu vou colocar aqui, é só ir clicando aqui nos pontos. É, vê se... Alice:

Ah, apareceu um join. Ih, cliquei em outra coisa e sumiu.

Entrevistadora: Ai, meu Deus. Vê se apareceu. Opa, peraí, deixa eu fazer de novo, então.

Renata: É, o meu sumiu também, gente. Eu cliquei em outra coisa. Pra variar.

Entrevistadora: Peraí que eu faço de novo aqui. É... Tá running, deixou terminado. E vou estar aqui. Aqui, só esse daqui e para cá.

Alice: Escolhe um, dois...

Entrevistadora: É, eu acho que assim, se vocês acharem que são todos, escolhe todos, mas se vocês acharem assim, sabe, um que não vale muito a pena...

Alice: Mas tem ordem, tipo assim, o primeiro a gente acha mais relevante. Dá para você entender que ele é o primeiro?

Entrevistadora: Não precisa, a ordem não vai ter relevância, não. Mariana: Lari, quando você diz mais relevante, é pensando no desenvolvimento de algoritmos...

Entrevistadora: Isso. Tá.

Mariana: E considera a diversidade.

Entrevistadora: Que impactariam, assim, coisas, questões até que para vocês, assim, foi interessante ver e falar, não, realmente isso é uma coisa que eu tenho que trazer mais para o meu dia a dia, para o desenvolvimento, para o pensamento, né, sistêmico do desenvolvimento.

Leandra: Acho que está limitado a três.

Entrevistadora: Está limitado a três? Peraí. Eita. Vou ter que fazer de novo, gente. Peraí. Eu tinha colocado, mas acho que ele fica no default. Foi mal.

Alice: Teve um spoiler ali.

Entrevistadora: Teve um spoiler mesmo. Poxa, que dó, vou jogar fora.

Renata: É legal esse negócio de... Mas acho que não vai mudar muito, não.

Entrevistadora: É, né. Peraí. Também acho que não.

Mariana: É, eu tinha escolhido dois.

Entrevistadora: Pronto, agora vocês podem botar em todos.

Mariana: Ih, mas não pegou tudo não, ele só está aparecendo para dois, três, quatro.

Alice: Agora que deu ruim de vez.

Regina: Mandar no chat.

Entrevistadora: É, né, eu acho que tecnologia que está... Quer dizer, o que está entre é tecnologia.

Mariana: Será que está com acessibilidade?

Alice: Acho que foi assim, tenta de novo e se limitar para três.

Entrevistadora: Ó, peguei todos eles, tá vendo? O teclado e a cadeira. Exatamente, é isso que está acontecendo. Aí tem seis votos.

Mariana: Ok, dois minutos. Bora lá.

Entrevistadora: Agora foi. Foi. E aí, se vocês quiserem já ir falando, né, assim, daqueles pontos que eu falei, né, se vocês acharam alguma coisa repetitiva, dar para suprimir, melhorar a escrita, alguma coisa que foi bem legal, alguma coisa que foi mais ou menos.

Regina: Eu acho que pelo andar da conversa, acho que você percebe algumas coisas que acabaram ficando repetitivas na hora da gente construir

a resposta. Mas eu não sei especificar, assim, de falar, tira o cartãozinho tal, tira tal. Mas acho que com a conversa você vê que o que está, como a gente foi pontuando as coisas versus o que está escrito, acaba levando para o mesmo caminho.

Entrevistadora: Sim. Não, dá para ver com a linguagem a mais, né?

Leandra: Uma sugestão que eu daria era se de repente juntasse dois cartões, sabe? É, meio disperso também, acho. E abria para a conversa.

Alice: Acho que poderia ficar um pouco mais fluido. Bora.

Leandra: Achei maravilhoso essa conversa aqui, né, que as meninas colocaram.

Eu acabei não falando depois porque muita coisa as meninas já estavam falando, né, então justamente para não ficar repetitivo.

Alice: Ai, nem lembro o que eu votei, meu Deus. Vai na fé aí, Larissa. Pelo amor de Deus.

Entrevistadora: Mas, poxa, que legal que você gostou, Leandra. O mais importante é é ter o envolvimento de vocês mesmos e perceber que é uma coisa que, né, até assim, é um pouco de dor de cada uma também, né, e como vocês se encontram também, né, nessas questões.

Alice: Ah, eu achei muito interessante, Larissa. Eu acho que você está construindo uma coisa muito legal assim, sabe, muito benéfica para todo mundo, não só para a gente que está aqui, entendeu? Mas eu espero que tenha um artigo aí, tenho certeza que vai ter. Mas está de parabéns, sério mesmo.

Está muito interessante. Acho que você soube fazer, sim, as perguntas. Talvez na hora de responder eu tenha me atrapalhado um pouco porque... É pouco, Larissa.

Pensar um pouco para a parte de produto é desafiador, né? Então, meu Deus, parece que eu estou fazendo um Enem aqui, confesso. Nem tudo para mim foi natural de se pensar, mas isso também é interessante. Por isso que eu acho que suas perguntas estão tão legais, sabe, porque dão espaço para esse debate, assim, não só do que a gente já está acostumada a falar, talvez.

Entrevistadora: Muito bom.

Renata: Oh, Larissa, eu acho que tem um ponto que... Caraca, ainda estão respondendo esse negócio de que vai participar, misericórdia. Gente, só um parênteses, porque eu ainda estou numa thread lá no e-mail. Misericórdia. O que eu acho que você poderia fazer... Essa pergunta 2 ela poderia ser a primeira. Por quê? A partir dela as pessoas já vão fazer a interpretação baseada nela.

Porque a primeira pergunta é aquela. Pô, qual é o tema? Aí você já encadeia tudo em cima dela, então fica mais fácil para o pessoal entender

aonde é que você quer chegar. Porque, por exemplo, a primeira pergunta ela é meio genérica.

Só que a segunda, ela já guia completamente aonde é que você quer chegar. E você nem precisa explicar muito, porque a segunda pergunta ela já diz o que você quer. Então eu acho que você pode trocar a ordem e talvez condensar duas perguntas em uma.

Aí você vê como é que você pode fazer e qual que você conseguiria condensar. Porque, primeiro, você conseguiria fazer em um tempo menor e a primeira, você não precisa explicar muita coisa, porque a pessoa já saberia onde é que você quer chegar.

Alice: Mas, assim, eu já acho interessante a primeira pergunta estar ali como primeira, justamente para a gente refletir nas dificuldades que a gente teria sobre o tema, sabe? Então, depende muito do que você quer, sabe, Larissa? Mas é uma visão minha, assim, concordo, que às vezes parece que é a mesma coisa, mas talvez não seja se você levar para um outro grupo, entendeu?

Entrevistadora: É, verdade. Vai depender da experiência de cada um. É, a ideia foi também um pouco, assim, deixar abrangente depois e começar fechando, né, e fechando, assim. Mas, muito bom.

Gente, super obrigada, de verdade, assim, nossa, vocês me ajudaram muito, as contribuições maravilhosas. Eu fico muito feliz que tem tanta mulher maneira nessa área, assim, e isso me dá muita motivação para continuar com a pesquisa, tá? Obrigada.

Mariana: Parabéns, Larissa, você, pelo convite, pela reflexão.

Sucesso, a gente quer que você tenha muito sucesso no seu mestrado e que possa escrever um artigo, como a Alice falou, né?

Alice: É doutorado. Doutorado. Doutora, futura doutora. É verdade, gente. Chique, chique. Antigo livro.

Outro patamar. Grande coisa. Boa sorte, né?

Renata: Se eu conseguisse montar grupos, tá, que abrangessem... O artigo é muito importante, faz parte da sua construção, assim, do seu projeto de doutorado.

Mas se você conseguir despertar essa consciência nas pessoas e que comece o debate e que comecem os grupos e você movimentar o ambiente, minha amiga, eu sei que para a academia talvez não valha muito, mas para a sociedade vale.

Leandra: É verdade. É uma comunidade, né?

Renata: A academia, ela quer número, ela quer artigo, ela quer resultado, nota, mas eu acho que quando a gente começa a provocar o debate na sociedade, você começa a instigar, bota aquela pulguinha atrás da orelha, sabe? Puxa,

uma mulher começa a comentar com outra, aí começam os grupinhos, começa a mudar a percepção, quando as mulheres se juntam, o bicho começa a pegar.

Regina: Estefany é muito maromba, cara. Estefany, minha amiga.

Renata: É porque a gente usa esse termo muito aí, é meio acadêmico.

Entrevistadora: Acabou, beleza.

Alice: Larissa, eu posso sair do negócio de estação, do negócio lá do Miro? Não preciso fazer nada não, né?

Entrevistadora: É só fechar. Gente, muito, muito obrigada. Eu vou dando status aí da pesquisa e se tudo der certo eu mando o link para vocês verem a defesa. Tá bom. Tchau, tchau.

Todas: Tchau, obrigada. Boa sorte. Tchauzinho, gente. Tchau, tchau.