



**Leonardo Freitas Sayão**

**Uma abordagem de Ciência de Dados para análise do  
impacto do viés cognitivo de busca de risco em tomadas  
de decisão individuais envolvendo perdas financeiras**

**Dissertação de Mestrado**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-graduação em Engenharia de Produção, do Departamento de Engenharia Industrial da PUC-Rio.

Orientadora: Prof. Fernanda Araujo Baião Amorim

Rio de Janeiro  
abril de 2024



**Leonardo Freitas Sayão**

**Uma abordagem de Ciência de Dados para análise do  
impacto do viés cognitivo de busca de risco em tomadas  
de decisão individuais envolvendo perdas financeiras**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-graduação em Engenharia de Produção, do Departamento de Engenharia Industrial da PUC-Rio. Aprovada pela Comissão Examinadora abaixo:

**Prof. Fernanda Araujo Baião Amorim**

Orientadora

Departamento de Engenharia Industrial – PUC-Rio

**Prof. Renata Silva Souza Guizzardi**

UT

**Prof. Leonardo dos Santos Lourenço Bastos**

Departamento de Engenharia Industrial – PUC-Rio

Rio de Janeiro, 08 de abril de 2024.

Todos os direitos reservados.

### **Leonardo Freitas Sayão**

Graduou-se em Engenharia de Produção pelo IBMEC-RJ em 2021. Desenvolveu seu projeto de pesquisa na área de Gestão de Negócios, nos temas de Apoio à Tomada de Decisão, Economia Comportamental e Ciência de Dados. Participou de eventos da área de Engenharia de Produção e Ontologia com apresentação de artigos a nível nacional. É analista de inteligência de mercado no ramo securitário.

#### Ficha Catalográfica

Sayão, Leonardo Freitas

Uma abordagem de Ciência de Dados para análise do impacto do viés cognitivo de busca de risco em tomadas de decisão individuais envolvendo perdas financeiras / Leonardo Freitas Sayão ; orientadora: Fernanda Araujo Baião Amorim. – 2024.

80 f. : il. color. ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial, 2024.

Inclui bibliografia

1. Engenharia Industrial – Teses. 2. Viés cognitivo. 3. Ciência de dados. 4. Aprendizado de máquina. 5. Tomada de decisão. I. Amorim, Fernanda Araujo Baião. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Industrial. III. Título.

CDD: 658.5

## **Agradecimentos**

De certo que, para alcançar o final desse grande ciclo, recebi muito apoio ao longo desse caminho. Em primeiro lugar, agradeço a Deus pelo dom da vida, pela saúde e disposição a mim concedidos.

À minha orientadora, Fernanda Baião, por me guiar e aconselhar nesse processo, tornando essa jornada de aprendizado ainda mais rica.

Aos colegas de curso, professores e funcionários do Departamento de Engenharia Industrial, sempre muito solícitos e agregadores, em prol de tornar o ambiente de ensino agradável. Em especial, ao Thiago e Mateus do nosso grupo de pesquisa, pelos estímulos e ideias.

À PUC-Rio, pelos auxílios concedidos, que foram muito bem aproveitados e que plantaram uma semente em prol do meu desenvolvimento.

À minha esposa, Juliana, pelo suporte e companheirismo sempre dispensados. Aos meus pais, Carla e Guilherme, por sempre terem me incentivado aos estudos. Aos meus irmãos, Paulo e Matheus, por serem compreensivos à minha rotina. Aos demais familiares e amigos que se importaram e me encorajaram a chegar ao final do curso.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

## Resumo

Sayão, Leonardo Freitas; Amorim, Fernanda Araujo Baião. **Uma abordagem de Ciência de Dados para análise do impacto do viés cognitivo de busca de risco em tomadas de decisão individuais envolvendo perdas financeiras**. Rio de Janeiro, 2024. 80p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

O estudo da tomada de decisões tem ganhado cada vez mais importância, desde as concepções clássicas do homem econômico até os mais recentes conceitos da racionalidade limitada e dos vieses cognitivos. Ao longo do tempo, a crescente complexidade das decisões impulsionou o desenvolvimento de tecnologias como os Sistemas de Apoio à Decisão e Modelos Preditivos, destacando-se mais recentemente a incorporação de técnicas do campo da Inteligência Artificial, e mais precisamente de Aprendizado de Máquina, para melhorar a precisão e a eficiência das tomadas de decisão. Entretanto, por maior que tenham sido os benefícios proporcionados pelos avanços no apoio computacional, as decisões são, em última análise, tomadas por humanos. E, sendo uma tarefa essencialmente humana, a influência dos vieses cognitivos em tomadas de decisão são um desafio relevante e pouco explorado. Esses vieses podem ser decorrentes de diversos fatores, incluindo preferências individuais, influências externas e derivações cognitivas inconscientes. Apesar dos esforços da área da Economia Comportamental em identificar e modelar esses vieses, seu impacto em contextos de decisões monetárias ainda é limitado. Portanto, este trabalho propõe uma arquitetura baseada em fundamentos ontológicos para identificar e analisar o impacto de vieses cognitivos em cenários de alto risco de perdas monetárias. Através da aplicação de técnicas de Ciência de Dados e Aprendizado de Máquina, o objetivo é propor uma metodologia implementada em um artefato computacional, capaz de automaticamente identificar padrões de vieses cognitivos a partir de um histórico de registros de decisões,

gerando conhecimento sobre as preferências de risco dos tomadores de decisão e seus ganhos e perdas diante das suas escolhas. O viés específico explorado neste estudo é a Busca de Risco no domínio de perdas, conforme definido no Padrão Quádruplo do Kahneman. A avaliação da eficácia dessa proposta será realizada por meio de um estudo de caso utilizando um *benchmark* disponível na literatura, fornecendo *insights* sobre a aplicabilidade e os benefícios práticos da arquitetura proposta.

## **Palavras-chave**

Viés cognitivo; ciência de dados; aprendizado de máquina; tomada de decisão.

## Abstract

Sayão, Leonardo Freitas; Amorim, Fernanda Araujo Baião (Advisor). **A Data Science approach to analyzing the impact of cognitive risk-seeking bias on individual decision-making involving financial losses.** Rio de Janeiro, 2024. 80p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

The study of decision-making has gained more and more importance, from the classical conceptions of the economic man to the more recent concepts of bounded rationality and cognitive biases. Over time, the increasing complexity of decisions has driven the development of technologies such as Decision Support Systems and Predictive Models, highlighting more recently the incorporation of techniques from the field of Artificial Intelligence, and more precisely Machine Learning, to improve the accuracy and efficiency of decision-making. However, as great as the benefits provided by advances in computer support have been, humans are ultimately the ones to make decisions. And, being an essentially human task, the influence of cognitive biases on decision-making is a relevant and underexplored challenge. These biases can be due to various factors, including individual preferences, external influences, and unconscious cognitive derivations. Despite the efforts of the field of Behavioral Economics to identify and model these biases, their impact in contexts of monetary decisions is still limited. Therefore, this work proposes an architecture based on ontological foundations to identify and analyze cognitive biases in scenarios of high risk of monetary losses. Through the application of Data Science and Machine Learning techniques, we propose a methodology - implemented in a computational artifact - capable of automatically identifying patterns of cognitive biases from a history of decision records, generating knowledge about the risk preferences of decision makers and their gains and losses caused by their choices. The specific bias explored in this study is Risk

Seeking in the loss domain, as defined in the Kahneman Quadruple Pattern. The evaluation of the effectiveness of this proposal will be carried out through a case study using a benchmark available in the literature, providing insights into the applicability and practical benefits of the proposed architecture.

## **Keywords**

Cognitive bias; data science; machine learning; decision-making.



## Sumário

<b>1.</b>	<b>Introdução</b>	<b>14</b>
1.1	Objetivos e Metodologia de Pesquisa	17
<b>2.</b>	<b>Referencial Teórico</b>	<b>20</b>
2.1	Tomada de decisão em um contexto de negócios	20
2.2	Vieses Cognitivos	21
2.3	Ontologia de Valor, Preferência, Risco e Viés	25
<b>2.4</b>	<b>Ciclo de implementação de Ciência de Dados</b>	<b>30</b>
2.4.1	Modelos de Aprendizado de Máquina	32
2.4.2	Métodos de avaliação de modelos de Classificação	38
<b>3.</b>	<b>Trabalhos Relacionados</b>	<b>40</b>
<b>4.</b>	<b>Uma arquitetura orientada a dados para análise de vieses cognitivos em tomadas de decisão</b>	<b>43</b>
<b>4.1</b>	<b>Aprendizagem de padrões dos dados por meios de técnicas de Ciência de Dados em um contexto consciente de viés cognitivo</b>	<b>46</b>
4.1.1	Cenário de Aplicação, Pré-processamento e Modelagem	47
4.2	Definição de Ganhos e Perdas para a Avaliação	48
<b>5.</b>	<b>Avaliação da proposta no cenário do <i>Choice Prediction Competition</i></b>	<b>51</b>
5.1	Projeto de Experimentação e Avaliação	51
5.2	Entendendo o cenário do CPC	52
5.3	Pré-processamento e transformação de dados	56
5.4	Experimento 1	60
5.5	Experimentos 2 e 3	63
<b>6.</b>	<b>Interpretações e Discussões</b>	<b>68</b>

<b>7. Conclusão</b>	<b>70</b>
<b>8. Referências Bibliográficas</b>	<b>73</b>
<b>9. Apêndice</b>	<b>80</b>

## Lista de Figuras

Figura 1 - O Padrão Quádruplo de Atitudes de Risco (Kahneman, 2011).	23
Figura 2 - Ontologia de Deliberação, Intuição e Decisão (Ramos et al., 2021).	26
Figura 3 - Deliberação, Intuição, Valor e Preferência na Ontologia de Decisão intuitiva (Ramos et al., 2021).	29
Figura 4 - Processo Knowledge Discovery in Databases (KDD).	30
Figura 5 – Processo CRISP-DM Data Science (Provost e Fawcett, 2016).	31
Figura 6 - Pseudocódigo resumo da operação do algoritmo Random Forest Classifier. Traduzido e adaptado de Hassan, et al. (2018).	34
Figura 7 - Pseudocódigo resumo da operação do algoritmo XGBoost.	35
Figura 8 - Pseudocódigo resumo da operação do algoritmo Regressão Logística. Traduzido e adaptado de Kumar (2021).	36
Figura 9 - Grid e Random Search de nove tentativas para otimizar uma função $f(x, y) = g(x) + h(y) \approx g(x)$ com baixa dimensionalidade efetiva (Bergstra e Bengio, 2012).	37
Figura 10 - Matriz de Confusão e métricas de avaliação.	38
Figura 11 - Arquitetura proposta para identificar e analisar vieses cognitivos no processo decisório.	45
Figura 12 - Fluxograma de pré-processamento dos dados.	56
Figura 13 - Fluxograma de criação de novas features (Feature Engineering).	58
Figura 14 - Acurácia dos Modelos por Cenários no Experimento 1.	61
Figura 15 - Matrizes de Confusão do Experimento 1.	63
Figura 16 - Acurácia dos Modelos por Cenário no Experimento 2 e 3.	65
Figura 17 - Matrizes de Confusão dos Experimentos 2 e 3.	67

## Lista de Tabelas

Tabela 1 - Hiperparâmetros a serem modificados na pesquisa no Random Forest.	34
Tabela 2 - Hiperparâmetros a serem modificados na pesquisa no XGBoost.	35
Tabela 3 - Hiperparâmetros a serem modificados na pesquisa na Regressão Logística.	36
Tabela 4 - Experimentos estabelecidos.	53
Tabela 5 - Amostra de dados do CPC (Erev, et al., 2017).	55
Tabela 6 - Games escolhidos do Choice Prediction Competition.	57
Tabela 7 - Variáveis excluídas da base.	59
Tabela 8 - Variáveis após o pré-processamento.	60
Tabela 9 - Variáveis explicativas do Experimento 1.	60
Tabela 10 - Melhores hiperparâmetros por Modelo do Experimento 1.	61
Tabela 11 - Acurácia por classe no Experimento 1.	62
Tabela 12 - Variáveis explicativas dos Experimentos 2 e 3.	64
Tabela 13 - Melhores hiperparâmetros por Modelo dos Experimento 2 e 3.	65
Tabela 14 - Acurácia por classe nos Experimentos 2 e 3.	66

# 1. Introdução

O fenômeno da tomada de decisões é estudado há muitos anos, Edwards (1954) relata uma extensa bibliografia sobre o assunto desde 1930. Um dos conceitos mais clássicos expostos por este autor são os do homem econômico, que possui as propriedades de ser completamente informado, dotado de sensibilidade infinita e um ser completamente racional. Essas foram as primeiras ideias a ganhar força e notoriedade no estudo das decisões na Economia e Psicologia. Anos mais tarde, essa ideia de racionalidade foi sendo questionada com base em experimentos que demonstravam que a decisão mais racional nem sempre era tomada, surgindo o conceito da racionalidade limitada (March, 1978).

Conforme os estudos foram avançando, a quantidade de decisões a serem tomadas aumentou cada vez mais em todas as áreas da sociedade. As novas tecnologias foram agregando valor nesse processo e encontrou-se a necessidade de auxiliar o tomador de decisão, de maneira a torná-lo mais assertivo e ágil.

Em tempos mais atuais, o auxílio ao decisor se deu de diversas formas, com destaque para Sistemas de Apoio à Decisão e Modelos Preditivos. Os Sistemas de Apoio à Decisão (*Decision Support Systems* - DSS) são sistemas de informação que ajudam as pessoas a usarem comunicações, dados, documentos, conhecimento e modelos por computador para resolver problemas e tomar decisões (Power, 2002). Os modelos preditivos podem prever o que irá acontecer sem que haja intervenção humana, mas com base em suas preferências, aumentando a confiabilidade e agilidade desses processos.

Técnicas de Inteligência Artificial (IA), e mais especificamente de Aprendizado de Máquina, vêm sendo incorporadas em DSS para melhorar cada vez mais as escolhas e para automatizar cada vez mais os processos oriundos dessas atividades. Essas abordagens observam diversos fatores, como: os atributos pessoais dos decisores, as características dos ambientes onde as decisões são tomadas, aspectos técnicos da elaboração de modelos e sistemas, dentre outros. Já existem estudos que correlacionam a atividade cerebral com a atitude de um indivíduo frente a uma decisão monetária (Engelmann, 2009). Por outro lado, muitos esforços têm sido feitos para automatizar decisões e criar modelos de aprendizado de máquina - como por exemplo, para maximizar o lucro na compra e venda de ações (Paiva et al., 2019).

Alguns vieses têm sido abordados no que tange à Ciência de Dados e à Tomada de Decisão. A título de exemplo, Mehrabi (2021) identifica e estuda dois tipos de vieses nesse contexto: o viés presente nos algoritmos de Aprendizado de Máquina e os vieses presentes nos dados disponíveis. No entanto, um tipo de viés de difícil aplicação que vem ganhando relevância são os vieses cognitivos (Tversky e Kahneman, 1991, 1992). Segundo os autores, esse viés emerge do Sistema Intuitivo de Decisão e pode comprometer as escolhas dos decisores (Kahneman e Tversky, 1979, 1984).

Das e Teng (1999) elencaram categorias principais em que os vieses cognitivos se incluem, estando ligados geralmente à exposição a alternativas limitadas, insensibilidade às probabilidades dos resultados, superestimativa da sua capacidade de gestão e o desenvolvimento de raciais concebidos anteriormente sobre os problemas de decisão que levem a conclusões inadequadas.

Segundo March (1993), os decisores muitas vezes criam análises e soluções que refletem suas experiências e as interpretam de forma a apoiar as suas crenças. Ou seja, por uma crença pessoal ou por uma influência externa, e muitas vezes de forma inconsciente, os vieses podem influenciar a tomada de decisão. Um exemplo é um tomador de decisão que esteja realizando uma perda em seus investimentos tende a mudar a sua posição para tentar voltar a realizar lucros (Thaler e Johnson, 1990). No entanto, isso pode ser tipicamente um caso de viés cognitivo, – como o Efeito *House Money*, definido por Thaler (1990), quando há aversão ao risco se houve perda e propensão a arriscar se houve ganho.

A despeito de seu impacto já ser reconhecido na literatura, os vieses cognitivos não têm recebido tanta atenção quando são estruturados *frameworks* para lidar com problemas de decisão. Seja pela vasta experiência ou elevada qualidade técnica do decisor, entre outros aspectos, todos os decisores humanos estão sujeitos a “armadilhas” cognitivas, quando o cérebro procura (muitas vezes inconscientemente) a alternativa mais viável e rápida que satisfaz um conjunto apresentado de restrições. Portanto, os vieses cognitivos costumam estar muito presentes durante um processo de tomada de decisão.

Apesar de parecer muito racional, decisões são comumente influenciadas por vieses cognitivos ainda desconhecidos pela maioria dos tomadores de decisão. Ainda que o campo da Economia Comportamental tenha muito a oferecer na identificação de vieses humanos na tomada de decisões, é difícil encontrar definições precisas e modelos de vieses adaptados a problemas específicos de negócios (Huyn et al, 2022), ou mesmo compreender seu impacto.

Portanto, o estudo dos vieses cognitivos dos indivíduos presentes nesses processos se destaca como possível ponto de melhoria das decisões. Muitos são os fatores que influenciam na incorrência de viés cognitivo – como as preferências individuais do decisor, o domínio em que a decisão está sendo tomada, o conhecimento técnico e os fatores externos.

Algumas aplicações foram feitas trazendo elementos da Economia Comportamental para dentro de Sistemas de Suporte à Decisão com o objetivo de tornar os argumentos de decisão mais reais e cognitivamente cientes de vieses. Arnott e Gao (2019) catalogaram os principais usos da EC nos DSS e mostraram que a principal estratégia de aplicação desses conceitos é realizando experimentos.

Usando esta estratégia, Ramos et al (2022) propuseram uma ferramenta de identificação automática do viés cognitivo de busca de risco com explicação durante a tomada de decisão envolvendo riscos e incertezas para apoiar decisores organizacionais. Uma pesquisa foi feita com decisores de organizações para confirmar a melhora da percepção dos participantes sobre suas preferências de risco. A detecção de vieses cognitivos por este trabalho, portanto, é realizada de forma pontual, a cada decisão tomada.

Tendo em vista que uma intervenção pontual gerou resultados, o histórico armazenado de todas as escolhas feitas dentro de uma organização – podendo ter uma intervenção ou não para a consciência de vieses – pode ser muito valioso para gerar novos *insights* para da área de negócio e desdobramentos do impacto dos vieses cognitivos nesse contexto, visando o aprimoramento das decisões tomadas.

Atualmente, muitos dados são produzidos a todo instante e as projeções futuras são ainda maiores. Isso leva ao cenário de ter grandes bases de dados com históricos sobre todo artefato computacional. Com o objetivo de auxiliar o tomador de decisão, é necessário estudar esse histórico para que a tomada de decisão seja mais ágil, correta e direcionada (Netto, 2021).

A crescente relevância dos dados como um recurso crucial para orientar decisões, impulsionada pelo aumento exponencial da disponibilidade de informações devido aos avanços tecnológicos, tem levado ao desenvolvimento da Ciência de Dados. Nesse contexto, uma abordagem essencial para extrair conhecimento estratégico é a aplicação de algoritmos e técnicas de Aprendizado de Máquina.

Na presente pesquisa, são integradas algumas práticas em um processo que engloba compreensão do problema, coleta e preparação de dados,

identificação de padrões, avaliação de modelos e implementação prática para ajudar em um problema de negócio (Provost, 2016).

Incluindo a presente introdução e preâmbulo (“Capítulo 1”), o presente trabalho está dividido em seis capítulos – cujos objetivos e metodologia da pesquisa serão especificados na sequência.

De forma específica, o Capítulo 2 traz o Referencial Teórico presente na pesquisa, compondo a teoria necessária para o desenvolvimento dessa pesquisa. O Capítulo 3 indica determinados Trabalhos Relacionados, seja pela metodologia aplicada, estratégias para mitigação de vieses cognitivos ou pelo domínio de atuação. No Capítulo 4, o objetivo é apresentar a arquitetura proposta por essa pesquisa de maneira geral, com diversos componentes que atuam em conjunto, com destaque para o módulo a ser detalhado no Capítulo 5. Adiante, o módulo de implementação de Ciência de Dados da arquitetura é especificado e detalhado para ser avaliado e instanciado no Capítulo 6.

## **1.1 Objetivos e Metodologia de Pesquisa**

O viés que será explorado nesse trabalho corresponde a um dos quadrantes do Padrão Quádruplo do Kahneman (2011), definido como Busca de Risco no domínio de perdas. Esse padrão é criado como forma de aproximar a ocorrência de determinados tipos de vieses nas decisões tomadas pelo sistema 1, ou sistema intuitivo.

A escolha do viés de busca de risco foi definida por possuir literatura consolidada, ter uma ontologia completa para delimitar seus conceitos, ser quantificável de maneira a trabalhar com algoritmos matemáticos e, por fim, por ter alta aplicabilidade em cenários reais de decisão.

A literatura e a ontologia serão detalhadas na Seção de Referencial Teórico. Esse tipo de viés tem uma natureza mais exata, definido por escolhas de probabilidades e valores. Assim, possui um melhor tratamento matemático onde podem ser aplicados modelos de aprendizado de máquina.

Sua aplicabilidade em cenários de decisão é vista dos cenários mais simples aos mais complexos. Muitos empreendedores passam por situações em que precisam decidir se irão realizar um prejuízo ou irão investir mais e tentam reverter o cenário desfavorável. Havendo ciência do decisor ou não, em muitos casos as chances de sucesso de reversão de um cenário de perdas são, geralmente, muito limitadas.

A opção mais arriscada e, portanto, passível de um viés de busca de risco é tentar reverter um prejuízo considerado certo. Por outro lado, realizar um



prejuízo menor no começo é a opção mais racional. Esse tipo de situação pode ocorrer com vários profissionais diferentes, com níveis de importância e valores distintos.

Dessa forma, a presente pesquisa tem o intuito de propor uma solução que possibilite o conhecimento do decisor sobre suas preferências de risco, levando em conta o histórico de decisões tomadas anteriormente dentro de um mesmo contexto, além de estruturar uma abordagem para entender o problema e os dados disponíveis, avaliando a ocorrência dos vieses cognitivos de busca de risco. Tendo em vista esse cenário e seus desdobramentos, a pergunta que norteou todo o desenvolvimento foi a seguinte: *“Em que circunstâncias o viés cognitivo de busca de risco contribui para a tomada de decisões mais valiosas?”*.

Sendo assim, os principais objetivos da pesquisa são:

- (i) Propor uma arquitetura, com fundamentos ontológicos, para identificar e analisar vieses cognitivos de busca de risco durante tomadas de decisão em cenários que envolvem riscos médios e altos de perdas monetárias;
- (ii) Desenvolver uma ferramenta computacional que implementa a arquitetura proposta, para identificar padrões de vieses cognitivos de busca de risco através de um processo baseado em Ciência de Dados com a aplicação de modelos de Aprendizado de Máquina, gerando conhecimento sobre os vieses existentes no histórico de decisões monetárias individuais envolvendo perdas monetárias.

A avaliação do atingimento de tais objetivos será feita a partir de estudo de caso usando a especificação de um benchmark disponível na literatura e que servirá de exemplo para a aplicação proposta.

Esta pesquisa, é de natureza majoritariamente quantitativa, mas com aspectos qualitativos em virtude da análise criteriosa dos fatores relativo aos vieses cognitivos e à base de dados.

Para atingir os objetivos delimitados, foram desenvolvidos dois artefatos principais: uma arquitetura de integração de um ambiente de dados e de tomada de decisão, e o detalhamento de operação de um módulo de ciência de dados para explicar os vieses cognitivos presentes no contexto.

A pesquisa bibliográfica foi feita no EBSCO e no Scholar, por uma string de busca com os termos de “Ciência de Dados” ou “Aprendizado de Máquina” ou “Inteligência Artificial”, “viés cognitivo” e “tomada de decisão”. Os trabalhos encontrados foram categorizados para verificar os campos estudo e as lacunas observadas.

A escolha do tema foi feita observando uma lacuna da literatura, de possuir uma abordagem de Ciência de Dados para tratar problemas de vieses cognitivos que possam ser quantificados em um contexto de tomada de decisão que envolva perdas financeiras. Isto é, de haver um tipo framework que oriente o uso de dados para explicar e analisar os vieses cognitivos inerentes à tomada de decisão.

O processo de Ciência de Dados foi implementado em linguagem de programação Python, com as principais bibliotecas de manipulação de dados e com as bibliotecas Scikit-Learn e XGBoost para aplicação dos modelos de Aprendizado de Máquina. Os processos incluíram a busca randomizada dos melhores hiperparâmetros e a validação cruzada dos modelos.

Os modelos foram escolhidos em virtude de terem funcionamentos distintos e por serem alguns dos modelos mais populares e interpretáveis. A Regressão Logística pode ser considerada como o modelo mais simples, que foi inserido para servir como uma linha de base para os demais. O *Random Forest* e o XGBoost foram selecionados porque são muito versáteis e tendem a dar bons resultados dada sua adaptabilidade (Kumari, 2017).

## **2. Referencial Teórico**

Este capítulo aborda todos os conceitos mais importantes para o embasamento e entendimento deste trabalho, atuando também como norteador para a abordagem a ser realizada. Dentro da temática da pesquisa, é crucial o entendimento sobre tomada de decisão, pois tudo deriva da necessidade de se tomar melhores decisões individualmente. Então, são indicados os conceitos mais importantes da Economia Comportamental, através dos vieses cognitivos, que dizem respeito ao mesmo contexto deste trabalho. Em seguida, será exposta a ontologia de risco, valor, preferência e viés, que formaliza uma visão estruturada dos conceitos e relacionamentos previstos em toda a teoria anteriormente mostrada, tornando a conceitualização sobre este tema mais explícita. Por fim, para o ciclo de vida de Ciência de Dados a ser conduzido na proposta, são apresentadas as justificativas para as etapas aplicadas na condução da pesquisa realizada, onde todos os preceitos da literatura serão colocados em prática, durante as seções subsequentes.

### **2.1 Tomada de decisão em um contexto de negócios**

Em um contexto de negócios, a escola mais reconhecida que estuda a tomada de decisão se baseia principalmente na teoria da racionalidade limitada (Simon, 1957). A tomada de decisão econômica é considerada racional quando maximiza essa utilidade esperada. A ideia do homem econômico é resumida pela lógica de que o perfeito tomador de decisões toma essas decisões com o conhecimento completo de todos os aspectos incluídos na situação, tem infinitas habilidades de processamento de informações, e não considera aspectos da psicologia humana.

De acordo com Simon (1957), a racionalidade dos tomadores de decisão era limitada e não perfeita. Ou seja, em vez de maximizar a utilidade, eles se acomodam nas circunstâncias e tomam a melhor decisão que podem em cada situação. Simon (1977) criou um modelo de fases para tomada de decisão contendo quatro etapas: inteligência, design, escolha e revisão. A teoria da racionalidade limitada explica que os tomadores de decisão usam heurísticas ou

regras básicas em vez de processos de otimização, o que aumenta o risco de tomar decisões enviesadas.

Existem ainda outras visões e teorias que buscam definir a tomada de decisão. Lindblom (1979) propôs uma estratégia aberta a ajustes conforme a necessidade, em uma lógica chamada de “confusão com um propósito”. Outra visão, mais direcionada para organizações e programas, busca manter o status quo dos procedimentos para cada situação. Com isso deve-se estudar cenários e pré-programar estratégias de ação mediante uma decisão para tomar (MARCH, 1962).

Cohen (1972) e Pfeffer (1981) contribuíram com outros direcionamentos para tomada de decisão. O primeiro representa o processo como uma lata de lixo, recebendo informações e influências de diversos tipos de agentes que agem de forma independente nas suas “micro decisões”, às quais estão sujeitos, sendo o produto de tudo isso o direcionamento global da organização. O último traz um aspecto político para os tomadores de decisão, para quando as decisões ocorrem em grupo, necessitam de negociações e são motivados por necessidades individuais, e não lógicas.

## **2.2 Vieses Cognitivos**

Em "*The Evolution of Cognitive Bias*" (Haselton, M.G. et al. 2015), o tópico é apresentado como um desafio para a psicologia evolutiva. Explica-se que para caracterizar esse viés, busca-se observar se a relação do traço cognitivo é de conformidade e continuidade com os processos de seleção natural. Dessa forma, comparam-se fatores de lógica e precisão com a resolução de um problema específico, ou seja, a proficiência e especificidade que o traço cognitivo pode trazer para identificar uma falha de projeto ou um bom uso de recursos psíquicos.

Haselton, M.G. (2015) tipifica os vieses cognitivos em três grupos: (i) Heurística, (ii) Gerenciamento de erros e (iii) Artefato. A primeira é a mais conhecida e estudada – trata do processamento limitado de informações pelo nosso cérebro, associando-o a restrições evolutivas. A segunda trata dos casos em que, por não haver uma solução específica para um problema, é utilizado um viés conhecido, mas que pode reduzir o erro em relação a outros métodos. A terceira é produto da colocação de humanos em ambientes não naturais ou da aplicação de padrões normativos inadequados, sendo tratados como aparentes estratégias de pesquisa.

Das e Teng (1999) fizeram uma revisão de literatura sobre vieses cognitivos, alocando em quatro categorias os que foram encontrados. São eles:

- a. Hipóteses anteriores e enfoque em alvos limitados. Trazer crenças ou orientações anteriores para o processo de análise e selecionam resultados e informações de interesse.
- b. Exposição a alternativas limitadas. Reduzir os problemas a racionais mais simples, procurando menos as alternativas disponíveis e confiando mais na intuição.
- c. Insensibilidade às probabilidades de resultados. Basear-se em julgamentos subjetivos em vez de características mais racionais do problema, tendendo a ver os problemas como únicos acabando por não levar em consideração as realizações passadas.
- d. Ilusão de capacidade de gestão. Tendência a ser excessivamente otimista, superestimando seu nível de controle da situação e acreditando que poderá gerir e corrigir suas ações ao longo do caminho.

A presente pesquisa busca tratar um viés específico que se encaixa melhor na primeira categoria definida por Haselton (2015) e na terceira categoria definida por Das (1999). O viés é o de busca de risco, segundo a definição apresentada pelo Kahneman (2011). Os vieses apresentados na Figura 1 compõem a *Prospect Theory* e estão orientados em um processo decisório onde as opções de escolha podem ser de alta ou baixa probabilidade de ocorrência e se tratando de valores positivos ou negativos. A busca de risco a ser aprofundada é a que se orienta no quadrante das perdas, ou seja, o viés que será explorado é o do canto superior direito.

No quadrante das perdas, observa-se um comportamento de procura de risco quando existem elevadas probabilidades de ocorrência dessa perda. Se um indivíduo já sabe que há grandes chances de perder, ele passa a querer evitar esse cenário, estando disposto a correr mais riscos. No quadrante dos ganhos, a lógica se inverte, pois quando há grandes chances de acumular algum valor, evita-se o risco (Kahneman, 2011).

	GAINS	LOSSES
HIGH PROBABILITY Certainty effect	<b>Risk Averse</b> e.g. Bet on the sure thing Fear of disappointment Accept unfavorable settlement	<b>Risk Seeking</b> e.g. Gamble between bad options Hope to avoid loss Reject favorable settlement
LOW PROBABILITY Possibility effect	<b>Risk Seeking</b> e.g. Lottery Hope to large gain Reject favorable settlement	<b>Risk Averse</b> e.g. Insurance Fear of large loss Accept unfavorable settlement

Figura 1 - O Padrão Quádruplo de Atitudes de Risco (Kahneman, 2011).

À época, Arnott (2005) identificou 37 vieses cognitivos existentes na literatura, em geral. Dentre eles, os vieses decorrentes de processos de negócios mais usuais estão resumidos abaixo.

(i) Aversão ao Risco (*Risk Aversion*). Com base na teoria do prospecto de Kahneman & Tversky (1984), esse tipo de viés explica que uma pessoa busca riscos para perdas e é avessa ao risco em relação a ganhos. A curva de utilidade 'S' ilustra esse conceito, que ainda é usado hoje em dia. Além disso, muitos estudos que tentam medir a percepção de risco de cada pessoa concluem que ela pode divergir.

(ii) Aversão à Perda (*Loss Aversion*). Conforme demonstrado por Kahneman e Tversky (1979), as pessoas valorizam mais as perdas do que os ganhos equivalentes. Indica que as perdas afetam mais as pessoas, mesmo que tenham origem em contextos diferentes, considerando o ponto de partida em que estiveram presentes.

(iii) Viés do status quo (*Status quo bias*). O viés do status quo é derivado da aversão ao risco. Ocorre pela apreensão que as pessoas têm em relação às mudanças, mesmo que as chances de perdas sejam mínimas. Esse viés também é observado quando se trata de questões não monetárias, com uma análise mais subjetiva que valoriza a manutenção do status quo (Samuelson e Zeckhauser, 1988).

(iv) Efeito dotação (*Endowment effect*). Também chamado de efeito de propriedade, os pesquisadores notaram que as pessoas não estão dispostas a desistir de um item, mesmo quando oferecem outro de valor semelhante. Os bens estão associados a si próprio, tornando-se um atributo da pessoa (Kahneman et al., 1989).

(v) Viés de confirmação (*Confirmation bias*). É a tendência das pessoas de recordar seletivamente as informações de maneira que beneficie seu palpite sobre

a situação. Esse viés leva a uma confiança sustentada por uma predisposição existente para que uma opinião seja confirmada com a interpretação de dados ou notícias que sustentem a crença predominante (Lord et al., 1979).

(vi) Viés de ancoragem (*Anchoring bias*). A pessoa está “ancorada” em um determinado ponto de vista, de modo que quando uma nova informação está disponível, ela é interpretada a partir desse primeiro conceito. Portanto, não reage objetivamente a novas informações e não faz uma avaliação correta do objeto sob análise (Sherif et al., 1958).

(vii) Viés de excesso de confiança (*Overconfidence bias*). Está associado ao julgamento subjetivo do decisor sobre sua compreensão da situação. Quando é necessário dar alguma resposta diante de um novo cenário, o indivíduo com esse viés constrói sua opinião com base em seus processos anteriores e não considera adequadamente o novo cenário (Fischhoff et al., 1977).

(viii) Efeito de recência (*Recency effect*). Intimamente relacionado com a heurística da disponibilidade, as pessoas confiam excessivamente nas informações mais recentes, baseando nelas as suas opiniões. Portanto, quanto mais recente o evento, mais disponível ele fica na memória do decisor, levando-o a acreditar erroneamente que já enfrentou mais situações semelhantes no passado (Tversky e Kahneman, 1973).

(ix) Viés da falácia de Gamber/viés da mão quente (*Gamber's fallacy bias/hot hand bias*). Com aplicação direta no mercado financeiro, esse viés mostra que as pessoas tomam retornos financeiros passados como base para esperar algum retorno específico no futuro, mesmo que os eventos não sejam correlacionados (Tversky e Kahneman, 1971).

(x) Efeito House Money. Thaler & Johnson (1990) propuseram, após os primeiros resultados sobre o que afeta a escolha do risco, que uma vez que o indivíduo teve um ganho anterior ou uma perda anterior, ele sofre desse tipo de viés: aversão ao risco se ele perdeu e propensão arriscar se você ganhou.

(xi) Viés retrospectivo (*Hindsight bias*). Traz a percepção do tomador de decisão sobre eventos passados. Conhecer o passado aumenta a crença de que esse evento poderia ter sido previsto (Fischhoff, 1975). As pessoas tendem a justificar seus erros e a se apropriar de situações de sucesso, mesmo que indevidamente (Roese e Olson, 2007). Isso pode levar a uma percepção errônea de acertos/erros anteriores, resultando em uma crença excessiva nas habilidades do tomador de decisão (Roese e Vohs, 2012).

(xii) Viés da escuta (*Herding bias*): Argumenta que mesmo havendo novas informações, elas podem ser ignoradas sob a justificativa de defender sua

reputação ou posição histórica sobre determinado assunto (Scharfstein & Stein, 1990). Aplica-se o “efeito manada” observado no mercado financeiro, tanto para gestores quanto para indivíduos (Barber et al., 2009; Lakonishok et al., 1992). Estudos apontam que a tendência de manada é mais forte em caso de mau desempenho passado (Merli & Roger, 2013).

### 2.3 Ontologia de Valor, Preferência, Risco e Viés

Ontologia pode ter significados diferentes, mas, no campo da Ciência da Computação, pode ser definida como uma especificação de uma conceitualização – ou seja, é uma descrição de conceitos e relacionamentos que existem entre estes (Gruber, 1995). É também trazida a ideia de ser uma especificação formal e explícita de uma conceitualização compartilhada (Studer et al., 1998).

Dessa forma, as ontologias são usadas para a base de conhecimento de um domínio e podem ser aninhadas em outras ontologias de maneira a enriquecer o campo de estudo (Noy e McGuinness, 2001). Então, através de uma ontologia, é muito mais fácil entender, organizar e estudar conceitos, estimulando seu compartilhamento e reprodução. Todos esses benefícios do uso das ontologias vão ao encontro dos objetivos dessa pesquisa, especificamente em dar mais destaque para os vieses cognitivos das decisões individuais.

Com o objetivo de suportar decisões estratégicas durante a definição do portfólio de projetos de Tecnologia da Informação (TI) em uma organização, Ramos (2022) colocou como seu alvo de análise os casos que se enquadram no *Risk Seeking* envolvendo perdas do Padrão Quádruplo do Kahneman (2011). Em uma situação típica citada, um projeto que vem recebendo investimentos e possui alta probabilidade de fracasso é alvo de uma decisão para o gestor: prosseguir com o projeto e aportar mais dinheiro, ou aceitar as perdas até então e encerrar o projeto. Essa decisão revela o apetite do risco do decisor: mesmo com grandes chances de perdas, ele pode arriscar confiando que seu projeto dará certo. Esse raciocínio é chamado de *Go/Kill decision*.

O autor desenvolveu a ferramenta ABI Tool, que identifica o viés de *risk seeking* em situações de incerteza, mas aplicado a decisões isoladas e em um contexto organizacional. A categorização das decisões tomadas (em decisões enviesadas ou não enviesadas) permite armazenar um histórico de todas as decisões tomadas na organização, incluindo a descrição do problema em que a decisão se contextualiza, as alternativas consideradas, cada uma com suas probabilidades de risco e possível impacto monetário, bem como a alternativa selecionada pelo decisor.



Guizzardi et al. (2020) defendem a importância de entender a natureza das decisões e o processo de tomada de decisão para fornecer melhor suporte à tomada de decisão e, por isso, propuseram a Ontologia de referência da Tomada de Decisão. Ramos et al. (2021) acrescentam à ontologia existente a tomada de decisão intuitiva de acordo com a *Cumulative Prospect Theory* (CPT) e introduzem conceitos como intuição e viés cognitivo (Figura 2).

Nessa ontologia, um agente (**Agent**) faz uma escolha (**Choice**) entre alternativas com uma certa intenção (**Intention**). A escolha é realizada por uma deliberação (**Deliberation**) fazendo uma escolha intuitiva (**Intuitive Choice**). Uma escolha intuitiva aplica uma intuição (**Intuition**), que, por sua vez, cria uma intenção denominada decisão (**Decision**). Ou seja, uma decisão é uma intenção criada por uma deliberação ou por uma escolha intuitiva. A escolha intuitiva pode ser influenciada por vieses cognitivos (**Cognitive Bias**), como busca de risco. Sendo uma intenção, essa decisão pode eventualmente ser manifestada através da realização de outra ação (**Action**) denominada ação resultante da decisão (**Decision Resulting Action**), cuja situação resultante é denominada consequência (**Consequence**).

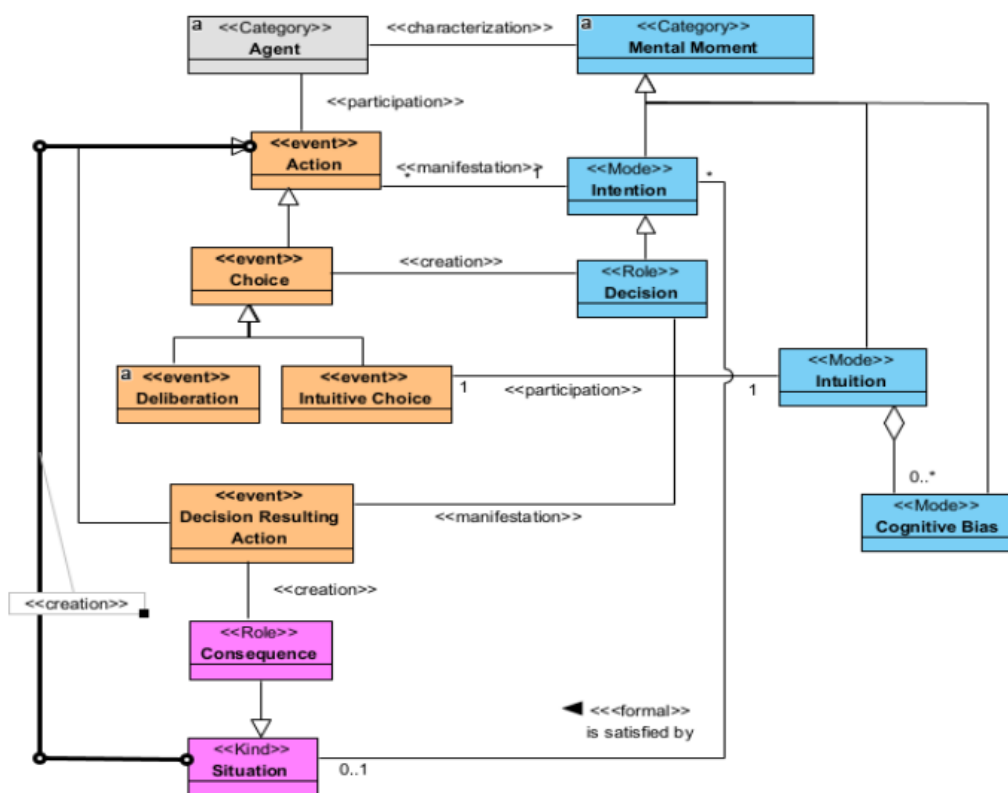


Figura 2 - Ontologia de Deliberação, Intuição e Decisão (Ramos et al., 2021)

Outra ontologia foi estendida e especializada para detalhar o Valor e Preferência de risco juntamente com a Deliberação e a Intuição, contemplando os conceitos de ponto de referência, aversão à perda, ganho, perda, aversão ao risco, busca pelo risco, atribuição de valor psicológico (Figura 3). A ontologia considera decisões sob risco e incerteza para tomada de decisão intuitiva e racional, e incorpora os conceitos de valor e preferência para decisões Go/Kill. Ainda, a ontologia proposta pelos autores serviu de base para codificar um esquema lógico de um banco de dados (BD) para um sistema de informação, que armazena de forma estruturada as informações relevantes para caracterizar tomadas de decisão e vieses cognitivos do tipo *risk seeking*.

Além da ontologia e do BD, os autores desenvolveram um método e uma ferramenta computacional de suporte ao método capazes de identificar vieses cognitivos (do tipo *risk seeking*) durante uma tomada de decisão intuitiva, aplicando as teorias de decisão racional (*Expected Utility Theory* e *Expected Value Theory*), a Teoria das Perspectivas (*Prospect Theory*) e as definições de vieses cognitivos de aversão ao risco (*risk aversion*) e de busca do risco (*risk seeking*) (LEWANDOWSKI, 2017).

A Figura 3 ilustra os processos racionais e intuitivos de tomada de decisão, bem como a relação entre os conceitos de valor e preferência. Quando um agente (**Agent**) precisa escolher entre duas alternativas, ele pode decidir de forma racional ou intuitiva, levando em conta suas próprias preferências (**Preferences**) em relação aos possíveis portadores de valor (**Value Bearer**). Cada alternativa é um portador de valor e o agente irá, por meio da deliberação ou da escolha intuitiva demonstrar suas preferências.

A tomada de decisão racional envolve uma deliberação (**Deliberation**) consciente, enquanto a intuição resulta em uma escolha intuitiva (**Intuitive Choice**). Utilizando qualquer uma dessas formas de pensar para fazer uma escolha, o agente atribui um valor à cada opção para decidir.

A atribuição de valor (**Value Ascription**) pode ser psicológica (**Psychological**) ou racional (**Rational**), dependendo se o agente utiliza o pensamento lógico (Sistema 2) ou intuitivo (Sistema 1). A atribuição de valor racional (**Rational Value Ascription RVA**) é resultado de uma avaliação deliberada das opções, enquanto a atribuição de valor psicológico (**Psychological Value Ascription PVA**) ocorre de maneira intuitiva e inconsciente. Ambos os sistemas operam em paralelo, sendo influenciados por componentes PVA e RVA (**Components**).

Os componentes de valor psicológico são afetados por vieses cognitivos (***Cognitive Bias***), como a aversão à perda (***Loss Aversion***), que influenciam a preferência do agente. Enquanto a Teoria da Utilidade Esperada presume que as pessoas escolhem otimizando, a atribuição de valor psicológico pode levar a uma preferência que difere daquela determinada pela Teoria da Utilidade Esperada, resultando em comportamentos como aversão ao risco para ganhos e busca ao risco para perdas.

O processo de tomada de decisão é descrito com base na deliberação racional ou escolha intuitiva, considerando as preferências do agente em relação aos portadores de valor. A decisão resultante pode ter implicações em como o agente avalia as consequências e influencia decisões futuras.

O autor também desenvolveu a ferramenta Automatic Bias Identification (***ABI Tool***), que identifica o viés de busca de risco em situações de incerteza em cada decisão isoladamente. A categorização das decisões tomadas (em decisões enviesadas ou não) permite armazenar um histórico de todas as decisões tomadas na organização, incluindo uma descrição do problema em que a decisão está contextualizada, as alternativas consideradas, cada uma com suas probabilidades de risco e possíveis prejuízos monetários, impacto e a alternativa selecionada pelo tomador de decisão.

Tal proposta, no entanto, não considera o histórico de decisões já tomadas como parte do conhecimento necessário para auxiliar a identificação automática de vieses cognitivos em tomadas de decisão futuras. Essa contribuição será utilizada diretamente na Seção 4, onde um esquema lógico de modelagem do banco de dados é criado para ser acessado por uma aplicação computacional de apoio à decisão.

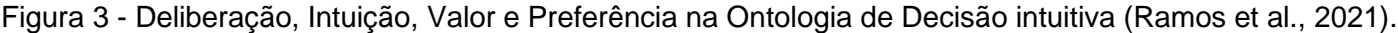


Figura 3 - Deliberação, Intuição, Valor e Preferência na Ontologia de Decisão intuitiva (Ramos et al., 2021).

## 2.4 Ciclo de implementação de Ciência de Dados

Historicamente, os dados sempre foram um recurso essencial para a tomada de decisões. Com os recentes avanços tecnológicos, o crescimento exponencial da quantidade de dados disponíveis trouxe às organizações um desafio de como extrair desses dados conhecimento útil. Nesse contexto, os processos de mineração de dados (*Data Mining*) surgiram como uma abordagem para extrair conhecimento estratégico de grandes coleções de dados, aplicando técnicas e algoritmos de Machine Learning. Mais recentemente, Nelli (2015) definiu Machine Learning como a disciplina que utiliza algoritmos e procedimentos para identificar padrões, tendências ou clusters, e a partir disso extrair informações valiosas para analisar dados de forma automatizada.

O processo de Ciência de Dados dessa pesquisa seguirá o padrão do Knowledge Discovery in Databases (KDD) (Frawley et al., 1992), conforme a Figura 5, a partir da base de dados que é o centro da arquitetura de integração de decisões, vieses e análises. Isso significa que a partir dos conceitos e relações lógicas de contexto escolhido os dados serão selecionados, pré-processados, transformados e submetidos a modelos de Machine Learning para identificar padrões e, por fim, gerar um conhecimento sobre esse domínio.

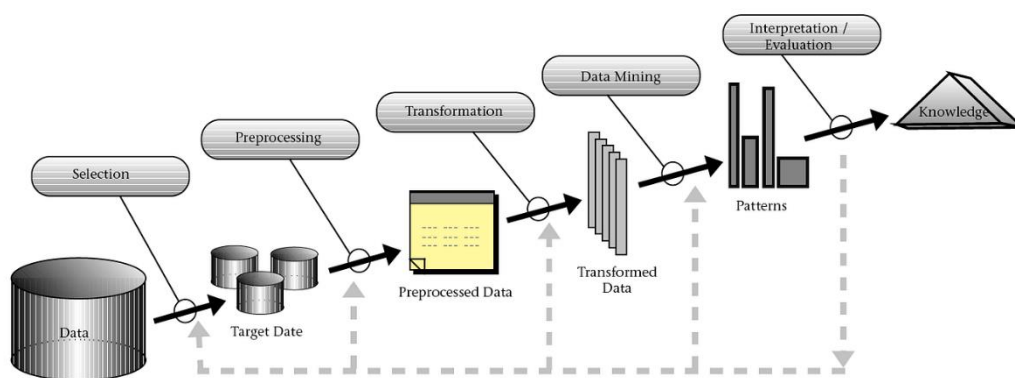


Figura 4 - Processo *Knowledge Discovery in Databases* (KDD).

Todas essas definições, metodologias e técnicas foram incorporadas à recente (e ainda crescente) área de conhecimento da Ciência de Dados (Priebe e Markus, 2015). Apesar de muitas mudanças terminológicas recentes neste tópico e áreas relacionadas, essa área é, essencialmente, um processo que engloba o entendimento do problema e definição da questão a ser investigada, a coleta e preparação de dados, a descoberta automática de padrões a partir dos dados,

avaliação dos modelos construídos com vista à questão de investigação e, posteriormente, implementação e utilização desse modelo em cenários reais.

Uma subárea da Ciência de Dados é a Mineração de Dados, representada na Figura 4 como a etapa central ao descobrimento de padrões (Frawley et al., 1992). Essa subárea se concentra na parte de modelagem e estatística para encontrar padrões nos dados, enquanto a Ciência de Dados engloba um processo maior e orientado à área de aplicação. A Figura 1 ilustra um processo padrão para iniciativas de Ciência de Dados, denominado CRISP-DM.

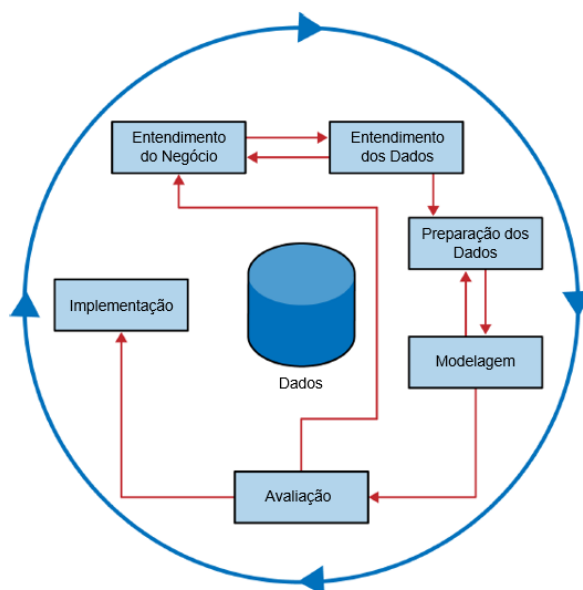


Figura 5 – Processo CRISP-DM Data Science (Provost e Fawcett, 2016).

Segundo o CRISP-DM, a compreensão dos dados e dos negócios possuem uma estreita ligação, e “se a solução do problema de negócios é o objetivo, os dados compreendem a matéria-prima disponível a partir da qual a solução será construída” (Provost e Fawcett, 2016). Em mesma proporção, enquanto aumenta o conhecimento dos dados melhor se torna a definição do ponto de vista do negócio. É como tomar uma decisão com maiores informações disponíveis, ao mesmo tempo em que as necessidades do negócio orientam a compreensão e exploração das bases disponíveis.

Na fase de preparação, ou pré-processamento, verifica-se a condição dos insumos disponíveis. Nessa etapa, percebe-se a qualidade dos dados, em termos de preenchimento dos requisitos, balanceamento de classes, dados faltantes e *outliers*, aplicando procedimentos de maneira a estruturar a base e resolver os problemas encontrados. Além disso, para a aplicação dos modelos, é essencial os dados estarem em formatos e estruturas adequados.

A modelagem dos dados consiste na aplicação dos algoritmos de Aprendizado de Máquina, cada um dentro de sua área atuação e com objetivos estabelecidos para atingir. Um modelo é escolhido de acordo com o problema apresentado e a pergunta que se busca responder a partir dele. Para cada tipo de problema há um modelo ou uma classe de modelos mais adequados. A Seção 2.4.1 tratará desse tema de maneira mais detalhada, apresentando o que é usado nessa pesquisa. Após a modelagem dos dados é necessário avaliar o ajuste aos dados, para, estatisticamente, ser possível tirar conclusões.

Na etapa de avaliação, os resultados dos modelos são analisados a fim de saber se geram previsões de boa qualidade, assim como medir a capacidade computacional requerida. A etapa de implantação (*deployment*) aplica os resultados da mineração de dados em uso real para resolver o problema inicialmente descrito.

#### 2.4.1 Modelos de Aprendizado de Máquina

Os Modelos de Aprendizado de Máquina são algoritmos desenvolvidos matemática e estatisticamente para aprender padrões a partir dos dados, sendo cada um desses modelos específicos para a resolução de uma determinada classe de problemas. O Aprendizado se divide de forma geral em três categorias: Supervisionado, Não Supervisionado e por Reforço. O primeiro desses será detalhadamente explicado a seguir, uma vez que é o ramo utilizado na pesquisa. As demais categorias serão apresentadas de forma objetiva na última parte dessa seção.

O Aprendizado Supervisionado é feito por algoritmos que necessitam de uma variável alvo para prever, isto é, buscam aprender os padrões que levam a uma determinada resposta para, então, reproduzi-las. Esse Aprendizado pode ser tanto numérico quanto categórico, e pode prever valores como em uma regressão ou uma categoria de dados em uma classificação. Essa pesquisa tem o objetivo de avaliar a classificação dos dados e a previsibilidade dos algoritmos escolhidos e, para tanto, buscou na literatura os modelos mais utilizados e adequados para tal.

Um grande *tradeoff* da escolha e ajuste dos modelos de Aprendizado de Máquina é conhecido como viés-variância, que ocorre devido à natureza dos algoritmos. Esse paradigma é o que dá origem ao *overfitting* e *underfitting* que são, respectivamente, um ajuste excessivo aos dados de treino ou um ajuste pouco sensível aos dados de treino. De maneira geral, quanto menos enviesado é o modelo maior é a sua variância e quanto menor a sua variância mais enviesado

o modelo se torna. O que se busca sempre com relação a esse dilema é o equilíbrio.

Uma técnica que tem sido utilizada para ajudar nessa questão são os métodos *Ensemble*, que combinam múltiplos modelos para produzir um único resultado. No caso de um problema de classificação, um método *Ensemble* irá gerar vários modelos e, com isso, é mais robusto em termos de viés-variância, produzindo previsões melhores.

Dois dos métodos *Ensemble* de Classificação mais utilizados são de *Bagging* e *Boosting*. No primeiro, busca-se reduzir a variância, evitando o *overfitting*, enquanto o segundo pretende reduzir o viés. Há ainda pelo menos uma grande diferença entre esses métodos que deve ser pontuada. No *Bagging* são gerados modelos paralelos e no *Boosting* os modelos são sequenciais para corrigir os erros anteriores.

Para agregar os resultados desses múltiplos modelos, o *Bagging* utiliza uma votação entre os vários modelos paralelamente criados enquanto o *Boosting* usa uma média ponderada para dar seu resultado. Dois exemplos de algoritmos que usam essas técnicas são o Random Forest e o XGBoost, que serão usados na Aplicação e Avaliação do cenário proposto, descritos na Seção 6.

Random Forest é um método de aprendizado de máquina baseado em árvores de decisão, cuja força reside na agregação de múltiplas árvores para mitigar o *overfitting* e melhorar a generalização do modelo (Breiman, 2001). Sua capacidade de lidar com dados heterogêneos e complexos o torna uma escolha sólida para diversas aplicações (Liaw, 2002).

O resumo da operação desse algoritmo está representado na Figura 5, contendo os principais passos aplicados nesse método, que incluem utilizar a Entropia e o Ganho de Informação para selecionar as características mais importantes dos dados e ramificá-los dessa forma (Hassan, et al. 2018).

<b>Random Forest Classifier</b>
<ol style="list-style-type: none"> <li>1. Selecione aleatoriamente <math>M</math> features do conjunto de features.</li> <li>2. Para cada <math>x</math> em <math>M</math>: <ol style="list-style-type: none"> <li>a. Calcule o Ganho de Informação: <math display="block">Gain(t, x) = E(t) - E(t, x)</math> <math display="block">E(t) = \sum_{i=1}^c c - P_i \log_2 P_i</math> <math display="block">E(t, x) = \sum_{c \in X} P(c) E(c)</math> </li> </ol> </li> </ol>



Where  $E(t)$  é a Entropia das duas classes,  $E(t, x)$  é a entropia da *feature*  $x$ .

- b. Selecione o nó  $d$  que tem o maior Ganho de Informação;
- c. Separe o nó em subnós;
- d. Repita os passos a, b e c para construir a árvore até atingir o número mínimo de amostras requeridas para separar;

3. Repita os passos 1 e 2 por  $N$  vezes para construir uma floresta de  $N$  árvores.

Figura 6 - Pseudocódigo resumo da operação do algoritmo Random Forest Classifier. Traduzido e adaptado de Hassan, et al. (2018).

Para um bom funcionamento do *Random Forest*, é necessário fazer uma boa escolha dos hiperparâmetros de inicialização do modelo. São diversas combinações possíveis e várias opções para serem ajustadas. Nessa pesquisa os hiperparâmetros a serem modificados a fim de produzirem as melhores previsões estão registrados na Tabela 1, juntamente com o significado de cada um, havendo ainda outros disponíveis na documentação do algoritmo para serem modificados.

#	Parâmetros	Significado
1	n_estimators	Número de árvores na floresta.
2	criterion	Função para medir a qualidade de separação da árvores, o Ganho de Informação.
3	max_depth	Profundidade máxima da árvore.
4	min_samples_split	Número mínimo de amostras requeridas para separar um nó interno.
5	min_samples_leaf	Número mínimo de amostras requeridas para ser uma folha do nó.
6	max_features	O número de <i>features</i> para considerar quando se considera o melhor <i>split</i> .

Tabela 1 - Hiperparâmetros a serem modificados na pesquisa no Random Forest.

O XGBoost, uma implementação otimizada de *gradient boosting*, destaca-se por sua eficiência computacional e desempenho superior. Projetado para lidar com grandes conjuntos de dados e dimensionalidade, segue uma abordagem de conjunto na qual vários modelos simples são treinados para produzir um modelo final mais robusto. O algoritmo constrói árvores de decisão a cada iteração, onde os modelos não são mais treinados de forma independente, mas sequencialmente, a partir do ajuste dos modelos previamente treinados (Chen, 2015).

<b>Algoritmo XGboost</b>	
<ol style="list-style-type: none"> <li>1. Inicializar <math>f_0(x)</math>;</li> <li>2. Para <math>k = 1, 2, \dots, M</math> faça: <ol style="list-style-type: none"> <li>a. Calcule <math>g_k = \frac{\partial L(y, f)}{\partial f}</math>;</li> <li>b. Calcule <math>h_k = \frac{\partial^2 L(y, f)}{\partial f^2}</math>;</li> <li>c. Determine a estrutura escolhendo divisões com ganho maximizado <math>A = \frac{1}{2} \left[ \frac{G_L^2}{H_L} + \frac{G_R^2}{H_R} - \frac{G^2}{H} \right]</math>;</li> <li>d. Determine o peso das folhas <math>w = -\frac{G}{H}</math></li> <li>e. Determine o algoritmo base <math>b(x) = \sum_{j=1}^T wI</math>;</li> <li>f. Adicione árvores <math>f_k(x) = f_{k-1}(x) + b(x)</math>;</li> </ol> </li> <li>Finaliza o loop.</li> <li>Onde <math>M</math> é o número de modelos base, <math>g_k</math> e <math>h_k</math> são modelos aditivos base, <math>L</math> é o risco a ser minimizado, <math>G</math> e <math>H</math> são a ponderação das folhas, <math>A</math> é o ganho a ser maximizado e os subscritos <math>L</math> e <math>R</math> são os lados esquerdo ou direito da árvore.</li> <li>3. Resultado: <math>f(x) = \sum_{k=0}^M f_k(x)</math>.</li> </ol>	

Figura 7 - Pseudocódigo resumo da operação do algoritmo XGBoost.

De modo semelhante ao *Random Forest*, o XGBoost também possui diversos hiperparâmetros que podem ser definidos antes da execução do modelo. A Tabela 2 elenca alguns dos principais que serão trabalhados nessa pesquisa e serão modificados nas execuções propostas com seus respectivos significados na implementação do modelo.

#	Parâmetros	Significado
1	learning_rate	É a redução do tamanho do passo usada na atualização para evitar <i>overfitting</i> .
2	n_estimators	Número de árvores na floresta.
3	colsample_bytree	É a proporção de subamostragem de colunas ao construir cada árvore.
4	max_depth	Profundidade máxima da árvore.
5	min_child_weight	Define a soma mínima dos pesos de todas as observações exigidas em um ramo.
6	subsample	Denota a fração de observações que são amostras aleatórias para cada árvore.

Tabela 2 - Hiperparâmetros a serem modificados na pesquisa no XGBoost.

A Regressão Logística, apesar de sua simplicidade, é uma ferramenta valiosa para modelagem preditiva quando a relação entre variáveis independentes e dependentes é de natureza logarítmica. Sua interpretabilidade e eficácia em problemas de classificação a tornam uma escolha clássica (Hosmer et al., 2013).

No detalhe da Figura 7, o algoritmo da Regressão Logística (Kumar, 2021) ao final atribui uma probabilidade de a classe ser 1 ou 0, em um problema binário. Isso passou a ser útil tanto nesse quanto tem outros modelos, para receber a probabilidade da classificação, podendo ser *input* para um outro modelo aperfeiçoando a previsão final. Por ser um modelo mais clássico, acaba por muitas vezes sendo escolhido como *benchmark* de classificação, mesmo sabendo que outros modelos podem performar melhor.

<b>Algoritmo Regressão Logística</b>	
1.	Inicializar $f_0(x)$ ;
2.	Para $i = 1, 2, \dots, K$ faça:
a.	Para cada instância de treinamento $d_i$ ;
b.	Defina o valor alvo para $z_i = \frac{y_i - P(1 d_j)}{[P(1 d_j)(1 - P(1 d_j))]}$ ;
c.	Inicialize o peso da instância $d_j$ para $[P(1 d_j)(1 - P(1 d_j))]$ ;
d.	Finalize a $f(j)$ para os dados com valor de classe ( $Z_j$ ) e o peso ( $w_j$ ) ;
e.	Decisão de classificação clássica;
f.	Atribua (class label: 1) se $P_{id} > 0.5$ , caso contrário (class label: 2);
	Finaliza o loop.
3.	Fim.

Figura 8 - Pseudocódigo resumo da operação do algoritmo Regressão Logística. Traduzido e adaptado de Kumar (2021).

A Regressão também possui alguns hiperparâmetros para ajustar, mostrados na Tabela 3 –como por exemplo a escolha do *solver* a ser aplicado na otimização e a penalidade usada, podendo ser Ridge, Lasso ou uma combinação das duas. O “C” também é de suma importância, pois dita se o modelo será mais ou menos complexo através da regularização mais fraca ou forte, evitando, assim, o *overfitting*.

#	Parâmetros	Significado
1	C	Inverso da força de regularização; quanto menor, maior forte a regularização.
2	penalty	Especificação da penalidade usada.
3	solver	Algoritmo a utilizar no problema de otimização.
4	fit_intercept	Especifica se uma constante deve ser adicionada à função de decisão.

Tabela 3 - Hiperparâmetros a serem modificados na pesquisa na Regressão Logística.

Além da seleção cuidadosa dos modelos, aprimorar o desempenho de cada algoritmo é crucial. Nesse sentido, foram empregadas estratégias avançadas de otimização de hiperparâmetros, utilizando o *Grid Search* e o *Bayesian Search*, visando alcançar a configuração mais eficaz para aumentar a acurácia de nossos modelos. Uma comparação entre os dois tipos de otimização pode ser vista na Figura 7, onde (a) é o *Grid* e (b) é uma *Random Search*. A busca Bayesiana é uma especificação da *Random*, então pode-se utilizar essa ilustração o entendimento da diferença entre a operação desses algoritmos.

O *Grid Search* é uma abordagem sistemática que explora manualmente um conjunto predefinido de valores para cada hiperparâmetro, avaliando o desempenho do modelo para cada combinação possível. Essa busca exaustiva permite identificar a configuração ideal, resultando em modelos mais robustos e precisos (Bergstra et al., 2012). Dessa forma, o gráfico da Figura 7 fica com seus parâmetros igualmente distribuídos, o que é definido na preparação do algoritmo.

O *Bayesian Search*, por sua vez, adota uma abordagem mais inteligente, incorporando informações acumuladas durante a busca para direcionar futuras tentativas. Essa estratégia bayesiana, baseada em probabilidades, permite uma exploração mais eficiente do espaço de hiperparâmetros, resultando em convergência mais rápida para configurações otimizadas (Snoek et al., 2012). Conforme a Figura 7, na letra (b), a otimização com aleatoriedade conseguiu atingir uma acurácia maior do modelo.

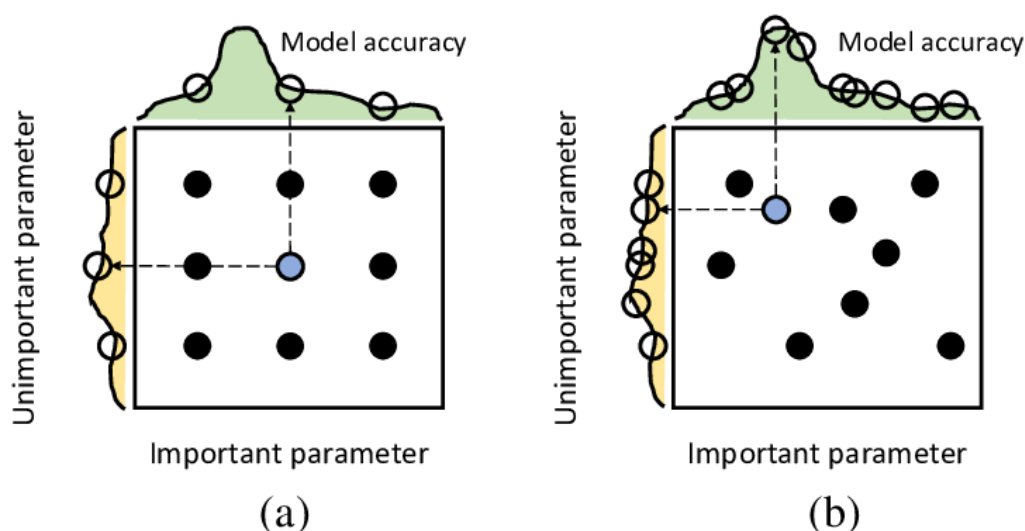


Figura 9 - Grid e Random Search de nove tentativas para otimizar uma função  $f(x, y) = g(x) + h(y) \approx g(x)$  com baixa dimensionalidade efetiva (Bergstra e Bengio, 2012).

### 2.4.2 Métodos de avaliação de modelos de Classificação

A avaliação de modelos de Aprendizado de Máquina é uma etapa crucial para compreender sua eficácia e desempenho em resolver tarefas específicas. Dentre as métricas utilizadas para essa finalidade, a matriz de confusão destaca-se como uma ferramenta fundamental (Provost, 2016). A matriz de confusão permite visualizar o desempenho de um modelo classificador ao comparar suas previsões com os valores reais dos dados de teste, representada na Figura 7.

Na matriz de confusão, os resultados do modelo são organizados em quatro categorias: verdadeiros positivos (TP), falsos positivos (FP), verdadeiros negativos (TN) e falsos negativos (FN). Com base nesses valores, diversas métricas podem ser derivadas para avaliar o desempenho do modelo, tais como precisão, recall, F1-Score e acurácia, todas elas também mostradas na Figura 7.

		Predito	
		Positivo	Negativo
Real	Positivo	Verdadeiro Positivo (TP)	Falso Negativo (FN)
	Negativo	Falso Positivo (FP)	Verdadeiro Negativo (TN)

$$\text{Precisão} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 * \text{precisão} * \text{recall}}{\text{precisão} + \text{recall}}$$

$$\text{Acurácia} = \frac{TP + TN}{TP + FN + TN + FP}$$

$$\text{Specificidade} = \frac{TN}{TN + FP}$$

Figura 10 - Matriz de Confusão e métricas de avaliação.

A precisão é a proporção de verdadeiros positivos em relação ao total de predições positivas, fornecendo uma medida da exatidão das previsões positivas. O *recall*, por sua vez, representa a proporção de verdadeiros positivos em relação ao total de casos positivos reais, medindo a capacidade do modelo em identificar corretamente todas as instâncias da classe positiva. O F1-Score é a média

harmônica entre precisão e *recall*, oferecendo um equilíbrio entre essas duas métricas. Já a acurácia é a proporção de predições corretas em relação ao total de predições, fornecendo uma visão geral do desempenho do modelo em todas as classes (Nelli, 2015).

Além dessas métricas, a matriz de confusão permite identificar padrões de erro do modelo, como a tendência a classificar erroneamente uma classe como outra (FP e FN), o que pode fornecer insights valiosos para ajustes e melhorias no modelo. De maneira geral, a matriz de confusão e as métricas derivadas dela desempenham um papel fundamental na avaliação de modelos de Aprendizado de Máquina, fornecendo uma visão abrangente do desempenho do modelo e orientando ajustes para otimização de sua eficácia (Raschka, 2019).

### 3. Trabalhos Relacionados

Existem na literatura abordagens de identificação de vieses cognitivos em tomadas de decisão, que serão resumidamente descritas nesta seção. O intuito é observar quais são as abordagens da literatura e quais as limitações no que tange analisar o impacto dos vieses cognitivos nas decisões tomadas. A busca por trabalhos sobre esses assuntos se baseou nos termos “vieses cognitivos”, “Ciência de Dados ou Aprendizado de Máquina ou Inteligência Artificial” e “tomada de decisão”.

No artigo de Huyn et al. (2022) o objetivo era aplicar um experimento, baseado em Economia Comportamental, para aumentar o volume negociações de uma loja, verificando, basicamente, qual melhor tipo de técnica de venda para cada público. Para tanto, foi feita uma pesquisa com clientes que foram à loja comprar produtos, em um caso real. Os tipos de clientes foram segmentados e os testes foram aplicados, em contraponto à linha de base (grupo controle). Várias abordagens foram definidas e catalogadas nos dados, de acordo com demografia e a característica do vendedor.

Todos esses dados foram submetidos a um modelo de Aprendizado por Reforço para que a política ótima fosse otimizada. O modelo foi usado para mostrar a melhor abordagem de venda e qual o melhor público e contexto para efetivar a venda. Foi utilizada simulação ABM (Agent Based Modeling) para testar a melhor política de venda para a loja com base nos resultados obtidos da modelagem anterior.

Analizando este trabalho com relação ao objetivo da presente pesquisa, ressalta-se que uma abordagem que leva em consideração aspectos e conceitos de Economia Comportamental deve se atentar as especificações do cenário, pois uma solução pode não ser eficaz em diferentes contextos.

Outro ponto que merece destaque é que, muitas vezes, não é possível testar os resultados em um cenário real e para tanto os autores utilizaram de uma Modelagem Baseada em Agentes para simular o comportamento dos consumidores. Embora tenham tido bons resultados com essas simulações, a ideia da presente pesquisa é que os dados reais sejam sempre utilizados, de

maneira que possam construir um histórico real sobre análise da tomada de decisão.

Bourgin et al. (2019) tem como objetivo apresentar um novo modelo de Aprendizado de Máquina pré-treinado para auxiliar nas aplicações de Ciência de Dados na área da Economia Comportamental. É dito que, a despeito da Economia Comportamental ter modelos teóricos premiados, a sua prática se desenvolveu lentamente e a construção de modelos preditivos robustos para decisão continua sendo desafiadora. O trabalho continua argumentando que a escassez de dados pode ser a principal responsável pela lentidão no avanço dessa área, uma vez que o comportamento humano requer tamanhos de amostra massivos para serem capturados com precisão pelos modelos disponíveis.

Para resolver este problema, os autores desenvolveram um modelo que foi pré-treinado por psicólogos cognitivos através de redes neurais, de maneira a ser menos suscetível a esse tipo de viés, testando o modelo em grandes conjuntos de dados e os autores denominam que os resultados alcançados foram “sem precedentes”. Apesar dos bons resultados e de nomearem o modelo como “Modelo Cognitivo”, não houve nenhuma especificação quanto a vieses cognitivos e quais aspectos cognitivos específicos esse modelo ajudaria. Para testá-lo, a sua aplicação é feita em grandes bases de dados, como a do *Choice Prediction Competition (CPC)*, que é um *benchmark* da literatura para a predição de decisões.

O CPC, por sua vez, foi uma competição focada em prever decisões para novos problemas com base em um histórico de dados robusto, com as descrições dos problemas e suas probabilidades (Erev et al., 2017; Plonsky et al., 2019). Apesar do grande foco nas decisões, não possui qualquer objetivo de explorar os vieses cognitivos desse contexto, o que, para o presente trabalho, foi visto como uma oportunidade de aplicação.

Existem outros trabalhos que exploram alguns tipos de vieses com Ciência de dados, mas não os vieses cognitivos. Mehrabi et al. (2021), por exemplo, pesquisam sobre preconceitos e injustiças em sistemas de inteligência artificial do ramo judiciário. Os vieses trazidos são de três grupos principais e estão presentes: (i) nos dados de entrada, (ii) nos algoritmos de Aprendizado de Máquina escolhidos, (iii) nos usuários que participam do processo.

O artigo explora com mais especificidade os preconceitos raciais e discriminatórios baseados na etnia. Os dados são estudados e tratados para que haja a aplicação mais livre possível desses tipos de vieses. Neste caso, o autor procurou abordar os três tipos de preconceito identificados ao mesmo tempo. O



terceiro grupo é o mais identificado com o presente trabalho, pois é na interação com as pessoas que surgem os vieses cognitivos.

Ramos (2021), através de uma abordagem fundamentada ontologicamente, cria uma ferramenta computacional capaz de alertar o decisor durante a tomada de decisão sobre seu posicionamento *risk seeking* – ou seja, enviesado cognitivamente para a busca do risco na sua escolha. Seu contexto é o de tomada de decisão envolvendo perdas monetárias de um portfólio de projetos. A abordagem bem construída e documentada pela ontologia torna facilitadas aplicações futuras. Embora o decisor seja alertado de sua posição de busca de risco, não há nenhuma aplicação considerando o histórico de decisões do mesmo para ser aproveitado e explorado no descobrimento de padrões.

As aplicações encontradas não relacionam diretamente os modelos de aprendizado de máquina com os vieses cognitivos dos indivíduos, para identificá-los e tratá-los durante um processo de tomada de decisão. O que se descobriu são técnicas para lidar com vieses nos próprios algoritmos, que, por simularem o pensamento humano ou por serem desenvolvidos por humanos, estão sujeitos a vieses cognitivos.

Os vieses cognitivos, por serem muito subjetivos, acabam ficando um pouco distantes do universo de dados em grande escala, pois as pesquisas geralmente são com tomadores de decisão de organizações ou empreendedores, sendo essa a lacuna da literatura que a presente pesquisa pretende contribuir.

#### **4. Uma arquitetura orientada a dados para análise de vieses cognitivos em tomadas de decisão**

Com base na fundamentação teórica exposta, na ontologia descrita no Capítulo 2 e nos objetivos do trabalho, propõe-se uma arquitetura que contemple ações em torno de um processo decisório consciente de vieses cognitivos e orientado a dados. Essa implementação computacional, que será descrita com mais detalhes no Capítulo 5, serve para obter conhecimento estatístico sobre os dados históricos, oferecendo predições de futuros casos.

A ideia central da arquitetura proposta é, a partir da ontologia exposta na seção anterior, derivar um ecossistema com suas relações lógicas definidas, facilitando a implementação e adaptação para casos semelhantes. Essa proposição estende o trabalho de Ramos (2021), que aborda o mesmo problema e viés que a ABI, para considerar as decisões tomadas historicamente e suas respectivas categorizações (com ou sem viés cognitivo) e implementar uma abordagem orientada a dados que identifica automaticamente o valor e as preferências de risco do usuário, para novas diferentes e posteriores decisões. Tal abordagem é implementada por um módulo de aprendizado de máquina capaz de identificar o valor e as preferências de risco do usuário em cada tomada de decisão.

A questão de investigação para este ecossistema é: *“Em que circunstâncias o viés cognitivo contribui para a tomada de decisões mais valiosas?”*.

Para endereçar essa questão, será desenvolvido um modelo de vieses cognitivos aprendido através de dados com uso de técnicas de aprendizado de máquina, para avaliar o impacto das decisões nos ganhos e perdas de cada decisor. Todos os componentes desta arquitetura podem ser vistos na Figura 4 e serão explicados com mais detalhes a seguir.

Na Figura 4, os retângulos em azul representam os módulos computacionais, que são os aspectos de um Sistema de Suporte à Decisão (DSS) que estão presentes nesse trabalho. Já os documentos em cinza representam dados de entrada/saída dos módulos computacionais e que são armazenados no banco de dados (cilindro cinza) para ação do módulo de Data Science e para

registrar todas as ocorrências de uma tomada de decisão. Os componentes estão em torno de uma base de dados que armazena o histórico de decisões tomadas, que serão insumos para os modelos de aprendizado de máquina descobrirem padrões ocultos e extrair conhecimento desses dados, conforme fluxo do KDD apresentado.

Pensando na aplicação prática dessa arquitetura, tomemos como exemplo um cenário real de uso para o seu melhor entendimento. Um decisor recebe um e-mail com diversas posições de sua carteira de investimentos que estejam dando prejuízo e deseja utilizar essa ferramenta computacional para auxiliá-lo nas escolhas. Então, os dados de entrada que constituem o problema de decisão estão representados pelo módulo Descrição do Cenário de Decisão, sendo os ativos que realizam prejuízo e suas opções de investimento para tentar recuperar o seu melhor valor. Esses dados são imputados através do Monitor de Decisão, e a partir de então os dados fazem parte de toda a cadeia de análise.

O usuário, então, toma sua decisão e registra, em cada uma das alternativas, se vai prosseguir ou não com os investimentos, que é o dado de saída Escolha da Decisão do Monitor de Decisão.

Nesse momento, entra em ação o Detector de Viés Cognitivo de Ramos et al (2021), que é responsável por notificar o decisor de suas decisões que contém o viés cognitivo de Busca de Risco. Em seguida, é gerado o alerta de Viés Detectado para o usuário e os registros são armazenados no BD. A Ontologia de Decisão Intuitiva representa a ontologia de tomada de decisão intuitiva e racional, sendo fundamental para a arquitetura. Através desta conceitualização é possível identificar e explicar os vieses cognitivos de busca de risco de acordo com a *Prospect Theory* (Kahneman, 1984).

O resultado de cada decisão, caso exista essa informação, é registrado através do Registro do Resultado das Decisões, que armazena esses dados no BD. O destaque desse módulo é para como esse registro de ganho ou perda pode ser feito. Cada indivíduo tem sua curva de utilidade de ganhos e perdas monetárias, então, a depender do decisor, um *outcome* pode ser mais ou menos intenso em utilidade para ele. Essa dualidade, entre o valor monetário e o valor percebido é também fundamentada nas teorias de Economia Comportamental.

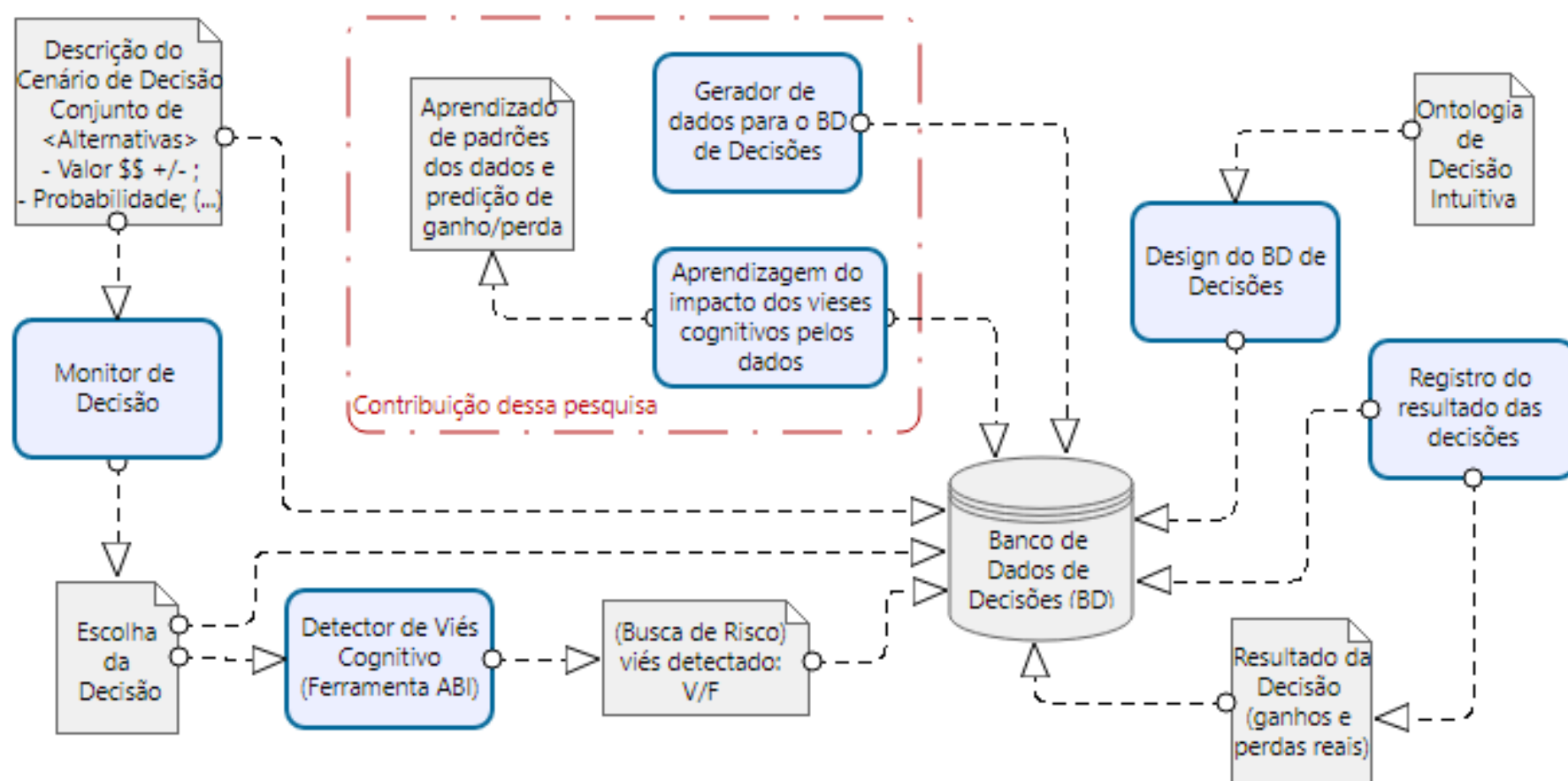


Figura 11 - Arquitetura proposta para identificar e analisar vieses cognitivos no processo decisório.

Essa conceitualização provida pela ontologia também foi utilizada como base no projeto do esquema lógico do BD, que é o Design do BD de Decisões. O módulo codifica o modelo conceitual baseado em ontologia em um modelo de banco de dados relacional para que os dados relacionados a cenários de decisão anteriores possam ser acessados, armazenados e transformados de forma rápida e segura em parâmetros computacionais.

O Gerador de Dados para o BD de Decisões e o Aprendizagem do impacto dos vieses cognitivos pelos dados são os principais componentes de contribuição deste trabalho. Juntos, eles atuam na parte *data-driven* para extrair padrões dos conjuntos de dados sobre o comportamento do tomador de decisão.

O Gerador de Dados para o BD de Decisões gera dados simulados ou coleta dados disponíveis publicamente para enriquecer o BD, como informações socioeconômicas sobre uma população ou indicadores geográficos de uma região.

O Aprendizagem do impacto dos vieses cognitivos pelos dados, uma das contribuições centrais d presente pesquisa, recebe todos os dados disponíveis, que forem de interesse, para aplicar algoritmos de aprendizado de máquina, com o objetivo de compreender os motivos que contribuíram para decisões tendenciosas. Esse módulo possibilita calcular os ganhos obtidos com os vieses, que é o dado de saída Aprendizado de padrões dos dados e predição de ganho/perda, possibilitando gerar a utilidade individual do decisor em sua percepção de como pode atingir os seus objetivos.

As saídas gerais da arquitetura são os resultados de algoritmos de aprendizado de máquina e a classificação de vieses durante a tomada de decisão. Todas essas informações são armazenadas no banco de dados e ficam disponíveis para consulta do sistema e utilização do modelo de ciência de dados. Outros desenvolvimentos estão sendo analisados e construídos baseados nesses resultados.

#### **4.1 Aprendizagem de padrões dos dados por meios de técnicas de Ciência de Dados em um contexto consciente de viés cognitivo**

Esta seção tem o objetivo de demonstrar o funcionamento do módulo de Aprendizagem do Impacto dos Vieses Cognitivos Através dos Dados, que foi apresentado na seção anterior, para posteriormente aplicá-lo em um cenário real. O módulo, presente na arquitetura, serve para atribuir melhores condições de decisão na medida em que explora os dados por meio das técnicas de

aprendizado de máquina, gerando o conhecimento de padrões antes desconhecidos e estabelecendo novas relações que podem ser úteis no estudo das decisões individuais.

#### 4.1.1 Cenário de Aplicação, Pré-processamento e Modelagem

Qualquer processo de Ciência de Dados, como o próprio nome mostra, precisa de dados para serem o insumo da análise. Na especificação proposta, uma base de dados compreende os registros de decisões tomadas por indivíduos, que essencialmente são as opções que podem ser escolhidas, com seus valores e probabilidades, e qual foi a escolha do indivíduo (Studer et al., 1998). Outras informações podem estar no registro e devem ser usadas para enriquecer a análise, como: idade, gênero, grau de escolaridade, cargo ocupado em uma empresa, área de atuação profissional, local de residência, dentre outros diversos fatores.

Esses dados “extra” irão depender da natureza da base de dados disposta – por exemplo, se for uma de clientes de um banco, possivelmente haverá informações de uso de cartão de crédito, salário, tipo de movimentações realizadas, dentre outras. O importante nesse estágio é ter pleno conhecimento dos dados em estudo, para que possam ser feitas as perguntas corretas, e para identificar assertivamente possíveis vieses cognitivos presentes. O entendimento do contexto e do problema de pesquisa é de suma importância para a análise dos dados (Priebe e Markus, 2015).

Conforme o desenho da arquitetura, apresentada na Seção 4, o módulo de Aprendizagem do impacto dos vieses cognitivos pelos dados possui apenas dois fluxos, sendo um deles o produto da sua implementação e o outro de onde provém todos os insumos para a análise. O BD um artefato que armazena os dados consultados por este desse módulo, para que possa haver um bom resultado. Diferentemente do módulo proposto, o BD possui oito fluxos que são direcionadas a ele, mostrando todo o tipo de informações que são armazenadas ali e que estão à disposição da aplicação do Aprendizado de Máquina.

A ontologia também serve como guia para a implementação de Bancos de Dados, conforme trabalho de Guizzardi (2008). Um BD bem estruturado é fundamental para que o módulo tenha uma boa operação. Os processos de criação, atualização, leitura e exclusão dos dados devem ser muito bem conectados aos demais módulos, para que não tenham dados faltantes ou imputados de maneira incorreta.

Para a Modelagem dos dados, ou seja, a aplicação dos algoritmos de Aprendizado de Máquina, deve-se observar qual a pergunta de pesquisa que irá orientar a escolha dos modelos a serem utilizados. Cada modelo, por sua vez, irá atender a um objetivo específico, sendo comum em um trabalho de Ciência de Dados a utilização de vários modelos (Mahesh, 2021).

Dada a sofisticação atual dos modelos, muito se estuda sobre a otimização e convergência para bons resultados, o que tem dado destaque para a seleção dos hiperparâmetros de cada modelos e suas interpretações, para extrair a melhor aplicação e predições possíveis para cada técnica e algoritmo (Raschka, 2019)

Assim, como cada modelo irá ajudar em um problema designado específico, cada tipo de técnica possui suas formas de avaliação, cabendo aos responsáveis pelo projeto a escolha das melhores métricas que façam sentido para a análise dos resultados e do contexto aplicado (Nelli, 2015).

## **4.2 Definição de Ganhos e Perdas para a Avaliação**

Em termos gerais, diz-se que uma decisão foi positiva ou negativa com base em seu resultado, ou seja, a posteriori. No caso de uma decisão monetária, o aspecto positivo está relacionado ao ganho de algum valor monetário, enquanto o negativo está relacionado à perda desse valor. Dessa forma, por muitos séculos, o cálculo do valor esperado (Whitworth, 1870) se tornou uma prática consagrada pelos matemáticos para fazer essas escolhas que se apresentam com valores e probabilidades.

No entanto, Bernoulli (Kahneman e Tversky, 1984) começou a mostrar que não era tão simples quanto parecia, já que cada indivíduo pode ter a sua percepção de ganho e perda particular gerada por sua função de utilidade. Em outras palavras, um mesmo valor monetário positivo ou negativo tem valor psicológico distinto para diferentes pessoas. Portanto, um ganho muito expressivo para um indivíduo pode ser menos relevante para outro e até mesmo pode representar um valor de perda.

Para não ter nem um ponto de vista 100% racional, com o valor esperado, nem 100% subjetivo com a utilidade esperada individual, os ganhos e perdas podem ser relativos, de acordo com um ponto de partida definido, sendo essa uma característica operacional do Sistema Intuitivo de Decisão (Kahneman, 2011).

Seguindo essa abordagem, a análise de ganhos e perdas dessa pesquisa foi feita com relação aos resultados de cada problema de decisão, sendo a diferença entre a alternativa escolhida e o resultado obtido de fato (registrado a posteriori) o que define se houve ganho ou perda.

Por exemplo, seja o cenário de investimentos em ações descrito na Introdução, no qual o portfólio do decisor está gerando uma queda de -5 unidades monetárias. O tomador de decisão deve escolher entre 2 alternativas sobre seu ativo:

- 1ª alternativa: realizar essa perda de 5, com certeza;
- 2ª alternativa: investir mais 5 unidades monetárias para seguir com o ativo aguardando uma valorização, que segundo seus estudos e projeções, possui 5% de chance de reverter a queda e dar 10 unidades monetárias de lucro. Neste caso, os 95% restantes dessa alternativa de escolha representam realizar uma perda de 10.

Suponha que o decisor tenha escolhido a segunda alternativa, e acabou realizando -10 unidades monetárias. Na arquitetura proposta, o decisor receberia duas informações:

- (i) um alerta de que sua decisão é enviesada por Busca de Risco, através da ferramenta *ABI Tool* descrita na Seção 2.3;
- (ii) uma predição de perda, com base no resultado do modelo de Aprendizado de Máquina (sabendo que essa perda seria de 5 unidades monetárias além das que já perderia na primeira alternativa).

Sendo assim, antes de confirmar a sua decisão, o decisor teria a oportunidade de revê-la (escolhendo a primeira alternativa para que não houvesse uma perda maior) ou prosseguir com a opção enviesada (mas tendo consciência de suas preferências de risco).

De uma maneira mais formal, essas são as possibilidades de ganhos e perdas nesse exemplo, com todas as opções de escolha:

- (i) Opção A = -5 e opção B = -10. Decisor escolheu A, logo seu ganho é de  $-5 - (-10) = +5$ . Ganho de 5 unidades em relação à B.
- (ii) Opção A = -5 e opção B = -10. Decisor escolheu B, logo sua perda é de  $-10 - (-5) = -5$ . Perda de 5 unidades com relação à A.
- (iii) Opção A = -5 e opção B = +10. Decisor escolheu A, logo sua perda é de  $-5 - (+10) = -15$ . Perda de 15 unidades com relação à B.
- (iv) Opção A = -5 e opção B = +10. Decisor escolheu B, logo seu ganho é de  $+10 - (-5) = +15$ . Ganho de 15 unidades com relação à A.

Para todos os problemas de decisão selecionados, a atribuição de ganho ou perda foi feita ao comparar o resultado da opção A com o resultado da opção B. No caso da opção A possuir um resultado maior que da opção B e o decisor a tiver escolhido, essa decisão foi rotulada como ganho. Do contrário, se uma opção



tiver como resultado um valor inferior e tiver sido escolhida, essa decisão foi rotulada como perda.

Nesse racional, como especificado nas próximas seções, não será considerado o valor absoluto do ganho ou da perda, sendo uma variável binária.

## 5. Avaliação da proposta no cenário do *Choice Prediction Competition*

Para fins de avaliação, o módulo de Ciência de Dados da arquitetura proposta será aplicado em um cenário *benchmark* de tomada de decisões existente, o *Choice Prediction Competition* (Erev et al., 2017). O objetivo é avaliar a capacidade de a arquitetura proposta gerar *insights* sobre o impacto dos vieses cognitivos a partir de um histórico de tomadas de decisão.

Todo o Ciclo de Ciência de Dados foi implementado através da linguagem Python. Os dados originais do CPC e os códigos utilizados nesse trabalho, bem como todos os resultados, estão disponíveis no GitHub<sup>1</sup>.

As próximas subseções descreverão o Projeto da Avaliação e Experimentação, a base de dados do CPC, o pré-processamento feito, os modelos aplicados e os resultados de cada um dos cenários de experimentação com suas respectivas análises.

### 5.1 Projeto de Experimentação e Avaliação

O Projeto de Experimentação e Avaliação da proposta avalia os objetivos desta pesquisa, que são:

- (i) Propor uma arquitetura, com fundamentos ontológicos, para identificar e analisar vieses cognitivos durante tomadas de decisão em cenários que envolvem riscos médios e altos de perdas monetárias;
- (ii) Desenvolver uma ferramenta computacional que implementa a arquitetura proposta, para identificar padrões de vieses cognitivos através de um processo baseado em Ciência de Dados com a aplicação de modelos de Aprendizado de Máquina, gerando conhecimento sobre os vieses de busca de risco existentes no histórico de decisões monetárias individuais envolvendo perdas monetárias.

Com relação ao primeiro objetivo, busca-se evidenciar a viabilidade da arquitetura proposta, e seu potencial para identificação de análise dos vieses cognitivos; já com relação ao segundo objetivo, a Experimentação visa prover resultados de avaliação sob três perspectivas:

---

<sup>1</sup> Github da pesquisa disponível em: <https://github.com/sayaoleo/BiasedDecisionsProject>

- 1) verificar como o atributo de viés cognitivo pode interatuar com os modelos de predição de ganhos e perdas;
  - 2) verificar como informações sobre o histórico das decisões de um mesmo indivíduo pode interatuar com os modelos propostos; e
  - 3) verificar se o *feedback* de uma decisão (*i.e.*, se o decisor souber que a decisão resultou em ganho ou perda) pode influenciar a predição do modelo quanto a ganho ou perda durante uma decisão.
- Para cada perspectiva, foi realizado 1 experimento.

A Tabela 4 sumariza a lógica dos experimentos realizados e as perguntas que buscam responder.

O Experimento 1 tem por objetivo verificar como o atributo de viés cognitivo pode interatuar com o modelo de predição proposto e, para isso, desenvolve modelos de predição de ganhos e perdas, comparando cenários com e sem o rótulo das opções de viés cognitivo de busca de risco.

Os Experimentos 2 e 3 têm por objetivo verificar como o histórico das decisões do decisor podem interatuar com o modelo de predição proposto e, para isso, desenvolve modelos de predição de ganhos e perdas utilizando informações sobre viés em decisões anteriores ou informações sobre o *feedback* recebido pelo decisor.

A próxima seção detalhará como foi feita a atribuição de valor de ganho e perda nos cenários a serem explorados.

## 5.2 Entendendo o cenário do CPC

A base de dados (dataset) de avaliação foi construída usando o conjunto de dados do *Choice Prediction Competition 18* (CPC 18) (Plonsky et al. 2017), contendo uma grande fonte de decisões tomadas por estudantes universitários israelenses sobre 270 problemas diferentes, os chamados “*games*”.

Essa competição consiste em um experimento com dados reais de opções de escolha em um jogo, com o objetivo de inferir decisões inéditas de indivíduos com base no histórico disponibilizado e prever a taxa média agregada de escolha de cada opção em cada jogo.

Os *games* envolvem valores monetários e probabilidades de sua ocorrência, sendo repetidos 25 vezes com cada respondentes. A partir da 6ª repetição ele recebe o *feedback* da sua decisão, ou seja, sabe o resultado da sua escolha (se ganhou ou perdeu). Quando o decisor recebe o *feedback* significa que ele teve acesso à quanto sua escolha retornou e quanto a escolha preterida retornou.

#	Experimentos	Perguntas	Objetivos intermediários	Objetivo geral
1	Desenvolver um modelo de predição de ganhos e perdas que possa ser utilizado com a informação de se a opção possui viés de busca de risco.	Se for considerada a identificação de viés cognitivo na decisão corrente, então o modelo de predição será mais bem interpretado e representará melhor os dados?	Verificar como o atributo de viés cognitivo pode interatuar com o modelo de predição proposto.	Criar um módulo de Ciência de Dados na arquitetura para prever ganhos e perdas nas decisões futuras utilizando vieses cognitivos como fonte de aprimoramento e interpretabilidade do modelo.
2	Desenvolver um modelo de predição de ganhos e perdas com a informação de haver ou não viés de busca de risco na última decisão tomada.	Se for adicionado um atributo de ocorrência de viés cognitivo na decisão anterior, então o modelo de predição será mais bem interpretado e representará melhor os dados?	Verificar como informações sobre o histórico das decisões do respondente pode interatuar com o modelo de predição proposto.	
3	Desenvolver um modelo de predição de ganhos e perdas que considere se o decisor teve viés em sua última resposta e se recebeu feedback da sua escolha.	Se for adicionado o atributo de Feedback no modelo, comparando decisões com e sem ele, então o modelo de predição será mais bem interpretado e representará melhor os dados?		

Tabela 4 - Experimentos estabelecidos.

Os tipos de *games* são muito variados, onde cada um tem sempre duas opções para decisão. Cada uma dessas opções pode ter até 10 valores possíveis para sua ocorrência, seguindo uma curva de probabilidade estabelecida. Os valores podem ser tanto de ganhos como de perdas e o respondente tem essas informações disponíveis no momento da escolha.

Esse cenário foi escolhido porque se encaixa para a aplicação da proposta, tendo todas as informações possíveis para explorar os quadrantes de Atitudes de Risco do Kahnemann (2011), com foco em um desses quadrantes (busca de risco envolvendo perdas) para uma correspondência com o classificador de vies da ABI *Tool*. A escolha do CPC como cenário de avaliação também considerou a existência na literatura de outras aplicações e estudos utilizando esses dados e o fato de ter seus dados disponibilizados gratuitamente, além de possuir uma grande variação nos problemas e nos respondentes.

O CPC foi utilizado como exemplo para estruturação de um possível ambiente de tomada de decisão, complementado por um conjunto de dados que eles disponibilizam como *benchmark* para registro dos resultados decorrentes das tomadas de decisão realizadas. A partir disso, a base de dados foi construída adaptando os dados do CPC à estrutura de dados fornecida pela arquitetura proposta.

Os dados foram obtidos diretamente do site do CPC<sup>2</sup>, onde são disponibilizadas diversas informações sobre o desafio, incluindo um artigo (Erev, 2017) para explicar melhor a origem do desafio, dos dados e dos objetivos da pesquisa. Um dicionário completo com todos os campos e explicações encontrados na base foi colocado no Apêndice deste trabalho, totalizando 30 colunas na tabela, conforme mostrado na Tabela 5.

A principal base utilizada é o chamado *Raw data*, que será usado na presente proposta e está exemplificado na Tabela 4. Essa base “*raw*” armazena todas as decisões tomadas, individualmente. Isto é, para cada linha da tabela há um decisor escolhendo sobre um problema diferente em uma tentativa diferente.

A Seção 6.2 descreve as etapas de pré-processamento dos dados que foram aplicadas.

---

<sup>2</sup> <https://cpc-18.com/>

SubjID	Location	Gender	Age	Set	Condition	GameID	Ha	pHa	La	LotShapeA	LotNumA	Hb	pHb	Lb	LotShapeB	LotNumBAmb	Corr	Order	Trial	Button	B	Payoff	Forgone	RT	Apay	Bpay	Feedback	block	
70109	Rehovot	M	35	7	ByProb	200	16	1	16	-	1	40	0.6	-15	-	1	0	0	22	1	L	1	40	16	80303	16	40	0	1
70066	Rehovot	M	25	7	ByProb	202	53	0.1	27	-	1	32	0.95	9	-	1	0	0	12	14	L	1	32	27	47240	27	32	1	3
61009	Technion	M	23	6	ByProb	160	60	0.1	19	-	1	34	0.8	23	-	1	0	-1	5	1	L	1	23	60	44318	60	23	0	1
60044	Rehovot	F	25	6	ByProb	171	-8	1	-8	-	1	7	0.25	-14	Symm	9	0	0	21	1	R	1	-14	-8	40587	-8	-14	0	1
70109	Rehovot	M	35	7	ByProb	205	36	0.05	16	L-skew	6	33	0.5	-1	-	1	0	0	4	1	L	0	16	33	37728	16	33	0	1
70081	Rehovot	M	27	7	ByProb	205	36	0.05	16	L-skew	6	33	0.5	-1	-	1	0	0	12	1	R	0	16	33	36784	16	33	0	1
70081	Rehovot	M	27	7	ByProb	203	104	0.05	1	Symm	7	34	0.05	2	L-skew	2	0	0	8	1	R	0	1	2	34737	1	2	0	1
70066	Rehovot	M	25	7	ByProb	202	53	0.1	27	-	1	32	0.95	9	-	1	0	0	12	5	L	1	32	27	32853	27	32	0	1
71051	Technion	M	22	7	ByProb	202	53	0.1	27	-	1	32	0.95	9	-	1	0	0	11	1	R	1	32	53	30743	53	32	0	1
61009	Technion	M	23	6	ByProb	153	39	0.4	6	Symm	5	15	0.95	12	-	1	0	1	12	1	L	0	39	15	28971	39	15	0	1
70109	Rehovot	M	35	7	ByProb	190	10	1	10	-	1	28	0.2	-1	-	1	0	0	10	1	R	0	10	28	28723	10	28	0	1
60058	Rehovot	F	25	6	ByProb	166	-6	1	-6	-	1	15	0.6	-10	-	1	0	0	23	1	R	1	-10	-6	27830	-6	-10	0	1
70109	Rehovot	M	35	7	ByProb	187	-3	1	-3	-	1	43	0.2	-20	-	1	0	0	28	1	L	0	-3	-20	26980	-3	-20	0	1
70095	Rehovot	F	23	7	ByProb	182	28	1	28	-	1	27	0.95	-11	-	1	0	0	2	3	L	1	-11	28	25818	28	-11	0	1
70095	Rehovot	F	23	7	ByProb	194	30	1	30	-	1	64	0.01	20	-	1	1	0	10	1	L	0	30	20	25816	30	20	0	1
70081	Rehovot	M	27	7	ByProb	195	10	1	10	-	1	92	0.01	9	Symm	9	0	0	15	1	R	1	9	10	23657	10	9	0	1
61006	Technion	M	23	6	ByProb	175	-4	1	-4	-	1	-2	0.8	-25	Symm	9	0	0	18	10	R	1	-1	-4	23586	-4	-1	1	2
70066	Rehovot	M	25	7	ByProb	210	25	0.2	10	-	1	102	0.05	9	Symm	3	0	0	29	1	R	1	9	10	23050	10	9	0	1
70081	Rehovot	M	27	7	ByProb	189	2	1	2	-	1	4	0.75	-23	-	1	0	0	14	1	L	0	2	-23	22777	2	-23	0	1
70095	Rehovot	F	23	7	ByProb	205	36	0.05	16	L-skew	6	33	0.5	-1	-	1	0	0	29	1	L	0	16	-1	22071	16	-1	0	1
60044	Rehovot	F	25	6	ByProb	154	46	0.1	16	-	1	65	0.2	2	Symm	5	1	0	16	1	R	1	2	16	21016	16	2	0	1
71051	Technion	M	22	7	ByProb	203	104	0.05	1	Symm	7	34	0.05	2	L-skew	2	0	0	10	1	L	0	1	2	19743	1	2	0	1
60034	Rehovot	M	28	6	ByProb	178	32	0.6	-38	-	1	14	0.2	4	L-skew	3	0	0	2	1	L	1	4	-38	19742	-38	4	0	1
70066	Rehovot	M	25	7	ByProb	202	53	0.1	27	-	1	32	0.95	9	-	1	0	0	12	1	L	1	32	27	18236	27	32	0	1
61075	Technion	M	30	6	ByProb	169	58	0.1	0	-	1	26	0.2	8	R-skew	3	1	-1	8	1	L	1	8	0	18090	0	8	0	1
60046	Rehovot	F	22	6	ByProb	176	-9	1	-9	-	1	12	0.5	-8	-	1	1	0	1	1	R	1	-8	-9	17857	-9	-8	0	1
70066	Rehovot	M	25	7	ByProb	207	22	0.9	-41	-	1	14	0.99	6	-	1	0	0	10	4	L	1	14	22	16801	22	14	0	1
61009	Technion	M	23	6	ByProb	179	36	0.8	-18	Symm	5	52	0.25	21	L-skew	3	0	0	1	1	L	1	52	35	16085	35	52	0	1
71070	Technion	M	20	7	ByProb	197	19	1	19	-	1	38	0.75	-33	R-skew	4	0	0	4	1	L	0	19	-33	15867	19	-33	0	1
61009	Technion	M	23	6	ByProb	164	1	0.5	1	-	1	40	0.01	-4	-	1	0	0	15	1	L	0	1	-4	15340	1	-4	0	1

Tabela 5 - Amostra de dados do CPC (Erev, et al., 2017).

### 5.3 Pré-processamento e transformação de dados

O fluxograma do pré-processamento aplicado ao cenário de avaliação é representado na Figura 12, relatando os passos desde a concepção desses dados até sua disponibilização aos modelos de Aprendizado de Máquina.

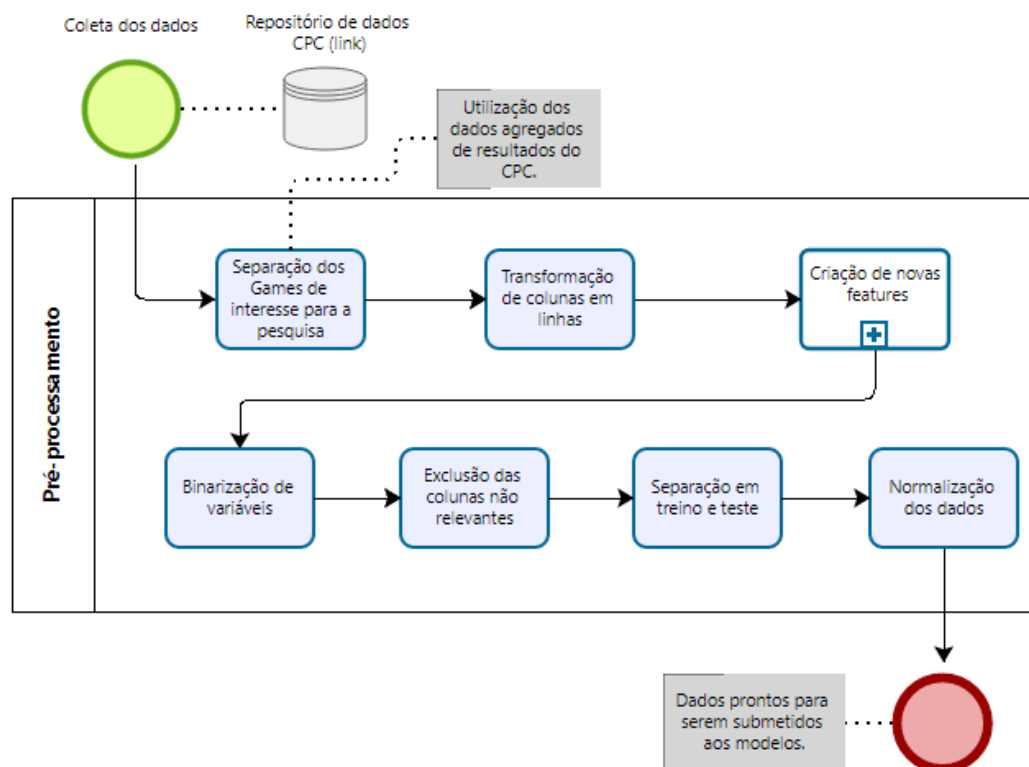


Figura 12 - Fluxograma de pré-processamento dos dados.

Considerando o cenário explorado nesse trabalho, de atender ao Quadrante do Padrão Quádruplo de Atitudes de Risco (Figura 1), apenas decisões que envolvem opções com altas probabilidades de perdas e com, no máximo, duas opções de escolha serão exploradas para uma correspondência com o classificador de viés da ABI Tool.

Dentre todos os tipos de games possíveis, a filtragem dos que atendem aos critérios estabelecidos pela pesquisa foi feita para respeitar esses limites de aplicação e para haver uma amostra de dados mais aderente aos objetivos. A base foi segregada para dispensar os tipos de decisões que não estão no domínio dessa aplicação, como por exemplo, decisões envolvendo ganhos (valores acima de zero).

Os dados foram, então, selecionados, chegando em um total de 8 games, mostrados na Tabela 6, correspondendo a 403 respondentes. Os games selecionados são os de número: 3, 5, 7, 33, 35, 42, 117 e 187, correspondendo a

um total de 20.375 linhas da base original, mas que ao final da presente etapa foram transformadas em 40.750 linhas, como produto do tratamento da base de dados. Ou seja, a base de dados utilizada em nossos experimentos continha um histórico de 20.375 problemas de decisão de 403 respondentes ao total. Como em cada decisão o tomador de decisão escolheu entre 2 alternativas, cada linha da tabela representa uma alternativa de decisão. Nas execuções dos experimentos, os dados foram selecionados nos *Trials* de 1 a 10, para que houvesse 5 tentativas sem o *feedback* e 5 tentativas com o *feedback*.

GameID	Opção A		Opção B			
	Valor	Probabilidade	Maior Valor	p Maior Valor	Menor Valor	p Menor Valor
3	-1	100%	0	50%	-2	50%
5	-3	100%	0	20%	-4	80%
7	-1	100%	0	95%	-20	5%
33	-3	100%	14	40%	-22	60%
35	-5	100%	47	1%	-15	99%
42	-6	100%	54	10%	-21	90%
117	-6	100%	7	50%	-30	50%
187	-3	100%	43	20%	-20	80%

Tabela 6 - Games escolhidos do *Choice Prediction Competition*.

A base do CPC possui um formato específico, conforme mostrado na Tabela 5. No entanto, o formato da base estabelecida na arquitetura (projetada a partir da ontologia descrita na Seção 2.3) é ligeiramente diferente. Para compatibilizar as estruturas de dados, foi necessário realizar o pivoteamento de algumas colunas e a inclusão de uma coluna, indicando a escolha ou não da opção, sendo 1 para a opção escolhida e 0 para a não escolhida.

Para viabilizar a execução dos experimentos propostos, foi necessário realizar *feature engineering*, isto é, a partir da base de dados original, derivar novas variáveis a serem incorporadas ao *dataset* que tragam novas e diferentes características aos dados que ainda não estavam explícitas na base de dados, conforme ilustrado na Figura 12 e detalhado a seguir.



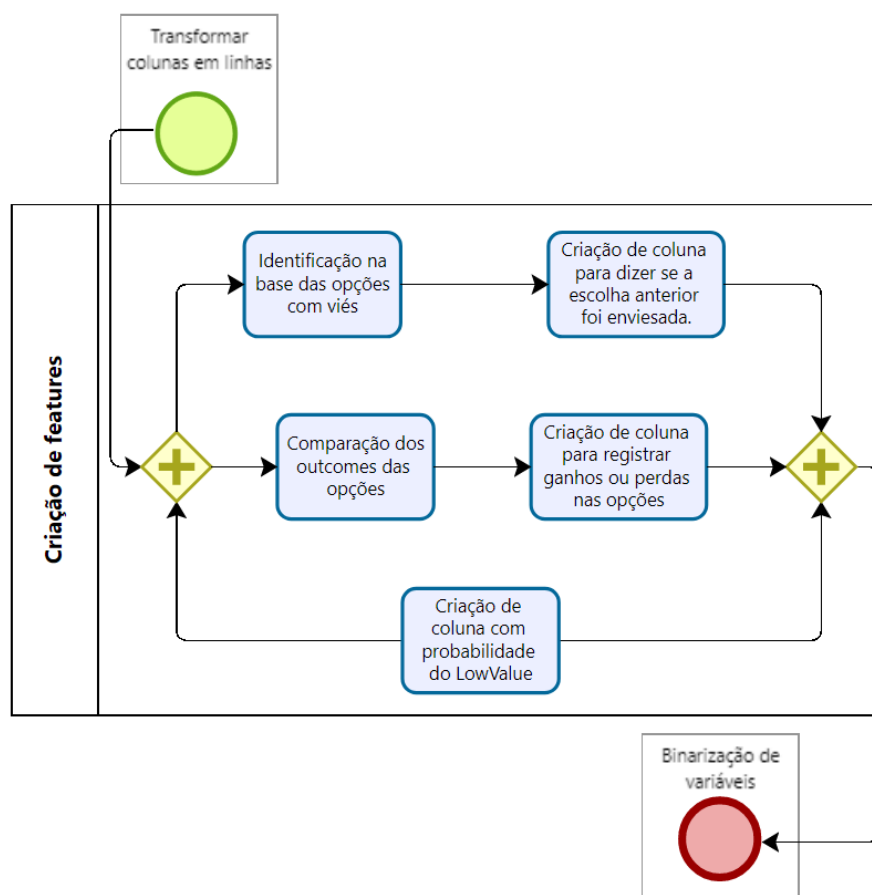


Figura 13 - Fluxograma de criação de novas features (*Feature Engineering*).

O principal objetivo da etapa de criação de novas *features*, detalhado na Figura 13, é trazer para a base as características relacionadas ao viés cognitivo de busca de risco, que não estão no experimento original do CPC, e medir a percepção de ganho ou perda com base nas opções disponíveis e resultados das escolhas. São originadas 4 novas colunas:

- (i) *BiasedOption* – se a opção possui viés de *risk seeking* ou não, sendo utilizada a ferramenta ABI. As opções foram classificadas com base no valor esperado das probabilidades.
- (ii) *RiskSeeking\_before* – se a opção anterior escolhida pelo respondente foi enviesada. A criação dessa coluna tem o objetivo de medir se saber se a última opção teve viés pode impactar o modelo de predição.
- (iii) *Gain/Loss* – relata se comparativamente à outra opção, houve perda ou ganho na escolha do decisor. Essa coluna poderá indicar se os vieses atuam de forma a aumentar os ganhos ou as perdas.
- (iv) *pLowValue* – apenas um ajuste para se adequar ao formato da base originada pela ontologia. A base original do CPC não possuía

essa probabilidade do menor valor, sendo feito apenas com uma subtração da probabilidade do maior valor, uma vez que são apenas duas opções.

A etapa seguinte no fluxograma da Figura 13 corresponde à binarização das variáveis categóricas de gênero (0 para masculino e 1 para feminino) e de ganho/perda das decisões (1 para ganhos e zero para perdas). Essa transformação é necessária para adequar os tipos de dados ao formato compatível com as técnicas de aprendizado de máquina aplicadas.

Os dados são constituídos de 41,5% de escolhas com viés de busca de risco, sendo que quem escolheu a opção enviesada na decisão corrente em 70,0% das vezes também escolheu a opção enviesada na tentativa anterior. Com relação aos ganhos e perdas obtidos, em 43,1% da base houve perda relativa à opção preterida.

A base original possui 30 variáveis, mas nem todas são relevantes para os objetivos determinados nessa pesquisa. Então, na Tabela 7 estão todas as variáveis e as que foram excluídas no começo desse processo, com sua justificativa. A maioria delas são atributos que não fazem sentido para a análise, como identificadores, dados ordinais e variáveis que são iguais para todos os registros da base selecionada.

#	Variável excluída	Motivo
1	SubjID	Identificador do respondente da pesquisa.
2	Location	Faculdade do respondente da pesquisa.
3	Set	ID do set que o respondente o fez.
4	Condition	Refere-se à ordem que o respondente viu os games.
5	GameID	Identificador do game realizado.
6	LotShape	Aplicável apenas para problemas com curva de distribuição.
7	LotNum	Número de opções disponíveis no game.
8	Option	Variável oriunda da etapa de pré-processamento.
9	Amb	Se as opções são ambíguas, segundo critério dos autores.
10	Corr	Se as opções são correlacionadas.
11	Trial	Número da tentativa dentro do game.
12	Order	Ordem que os games foram jogados pelos respondentes.
13	Button	Lado da tela que o botão da resposta se localiza.
14	Choice	Variável oriunda da etapa de pré-processamento.
15	Payoff	Quanto a opção escolhida pagou.
16	Forgone	Quanto a opção preterida pagou.
17	RT	Tempo de reação do respondente no game.
18	Apay	Valor de resposta da opção A.
19	Bpay	Valor de resposta da opção B.
20	block	Divisão ordinal acima do Trial.

Tabela 7 - Variáveis excluídas da base.

Portanto, a Tabela 8 traz todas as variáveis que foram submetidas aos modelos de predição, com suas respectivas descrições. A última variável dessa tabela é a chamada “alvo”, pois será por meio dela que o modelo buscará aprender os ganhos e perdas para predição.

#	Variável	Descrição
1	Gender	Gênero do respondente, sendo 1 para masculino e 0 para feminino.
2	Age	Idade do respondente.
3	HighValue	Maior valor disponível no Game.
4	pHighValue	Probabilidade do maior valor disponível no Game.
5	LowValue	Menor valor disponível no Game.
6	Feedback	Indica se o Trial teve feedback da decisão, variável binária.
7	pLowValue	Probabilidade do menor valor disponível no Game.
8	BiasedOption	Atribui à opção corrente ser enviesada por busca de risco, variável binária.
9	RS_before	Indica se a última decisão do respondente foi enviesada.
10	G/L Sign	Variável binária para mostrar se o game teve ganho (1) ou perda (0).

Tabela 8 - Variáveis após o pré-processamento.

Os próximos capítulos irão detalhar os experimentos e seus resultados. Ressalta-se que para cada experimento delineado foram utilizadas as técnicas de randomização no processo de melhora dos hiperparâmetros dos modelos e validação cruzada nos dados de treino. Sendo os melhores resultados dos dados de treino aplicados para os dados de teste, que correspondem a 20% da massa total de dados.

## 5.4 Experimento 1

O Experimento 1, conforme a Tabela 4, busca verificar se o conhecimento de quais alternativas de decisão estão enviesadas cognitivamente interfere na capacidade do modelo de predizer se haverá ganho ou perda. Nesse experimento, o conjunto de dados do CPC recebe a incorporação da variável binária *BiasedOption*, que indica, para cada alternativa de decisão, se ela possui ou não viés de Busca de Risco. As variáveis usadas em cada Experimento estão elencadas na Tabela 9, sendo os dois Cenários comparados neste experimento:

- Cenário 1: sem a variável *BiasedOption*;
- Cenário 2: com a variável *BiasedOption*.

Cenário / Variáveis	HighValue	pHighValue	LowValue	pLowValue	BiasedOption
Cenário 1	X	X	X	X	
Cenário 2	X	X	X	X	X

Tabela 9 - Variáveis explicativas do Experimento 1.

Os algoritmos de classificação aplicados foram o *Random Forest*, XGBoost e Regressão Logística, conforme descritos na Seção 2.4.1. Para cada algoritmo, foi executada uma rotina de varredura de hiperparâmetros via Bayesian Search (Snoek et al., 2012), buscando a melhor configuração de cada algoritmo. As melhores configurações encontradas estão descritas na Tabela 10.

Modelos	Melhores Hiperparâmetros	
	Cenário 1	Cenário 2
Random Forest	criterion: gini, max_depth: 2, n_estimators: 37, max_features: 27, min_samples_leaf: 34, min_samples_split: 3	criterion: log_loss, max_depth: 2, n_estimators: 90, max_features: 41, min_samples_leaf: 9, min_samples_split: 9
XGBoost	learning_rate: 0.1, max_depth: 4, n_estimators: 10, subsample: 0.75, colsample: 1, min_child_weight: 1	learning_rate: 0.1, max_depth: 8, n_estimators: 10, subsample: 0.75, colsample: 0.5, min_child_weight: 1
Logistic Regression	C: 0.0001, fit_intercept: False, penalty: l2, solver: lbfgs	C: 0.0001, fit_intercept: True, penalty: l2, solver: liblinear

Tabela 10 - Melhores hiperparâmetros por Modelo do Experimento 1.

Os melhores resultados dos modelos foram sumarizados na Figura 14, mostrando a melhor acurácia de cada uma das execuções. Nota-se que o modelo com melhor performance é o XGBoost, com uma acurácia de 78,88% e uma diferença de 2,93 pontos percentual para o Random Forest. A Regressão Logística, conforme esperado, teve um menor desempenho, com 70,57% de acurácia.

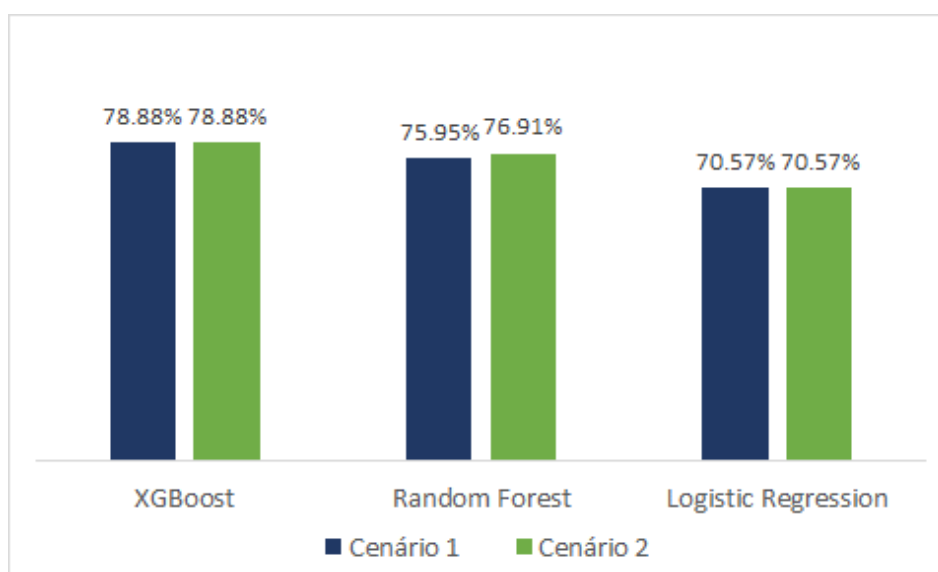


Figura 14 - Acurácia dos Modelos por Cenários no Experimento 1.

Em seguida, foram calculadas as acurácias por classe de cada modelo, a fim de avaliar se a capacidade de predição de cada modelo variou entre ganhos e perdas. Os resultados encontram-se na Tabela 11.

Cenários	Modelos	Acurácia			Matriz de Confusão				
		Classe=Perda	Classe=Ganho	Total	TP	FP	TN	FN	Total
Cenário 1	XGBOOST	74.23%	83.73%	78.88%	602	117	556	193	1468
	Random Forest	82.92%	72.51%	75.95%	712	270	403	83	1468
	Regressão Logística	72.61%	69.41%	70.57%	649	286	387	146	1468
Cenário 2	XGBOOST	74.23%	83.73%	78.88%	602	117	556	193	1468
	Random Forest	80.93%	74.57%	76.91%	692	236	437	103	1468
	Regressão Logística	72.61%	69.41%	70.57%	649	286	387	146	1468

Tabela 11 - Acurácia por classe no Experimento 1.

Conforme pode ser observado na Tabela 11, de maneira geral, o XGBoost obteve melhor acurácia na predição dos ganhos, com 83,73%, enquanto o Random Forest teve uma maior acurácia para as perdas, com 82,92%. Ou seja, o XGB teve melhor desempenho para predizer ganhos, enquanto o RF para predizer perdas. Em segundo plano, a Regressão Logística foi o modelo mais equilibrado, possuindo taxas mais similares de predições corretas tanto nas perdas quanto nos ganhos, de 72,61% e 69,41% respectivamente.

Ainda, as acurácias por classe não tiveram diferenças significativas entre os cenários, de onde se conclui que a informação sobre a existência de viés cognitivo não influenciou a capacidade de predição do modelo. Isto pode ser decorrente do fato de que todas as informações necessárias para rotular uma alternativa como sendo enviesada estarem disponíveis no *dataset*, tornando o modelo capaz de levar em consideração o viés cognitivo, mesmo sem ter conhecimento explícito sobre a existência dele.

Todas as Matrizes de Confusão referentes ao Experimento 1 encontram-se na Figura 15, onde é possível observar de maneira gráfica onde estão os maiores índices de acerto dentre os modelos e categorias. Quanto mais claras as cores, próximas ao tom do amarelo, maiores as concentrações dos casos nos quadrantes. Quanto mais escuras as cores, próximas ao tom do vinho/roxo, menores as proporções do quadrante com relação ao total de casos. Dessa forma, uma ordem geral de cores com relação aos melhores valores são: Amarelo > Verde > Azul > Roxo/Vinho.

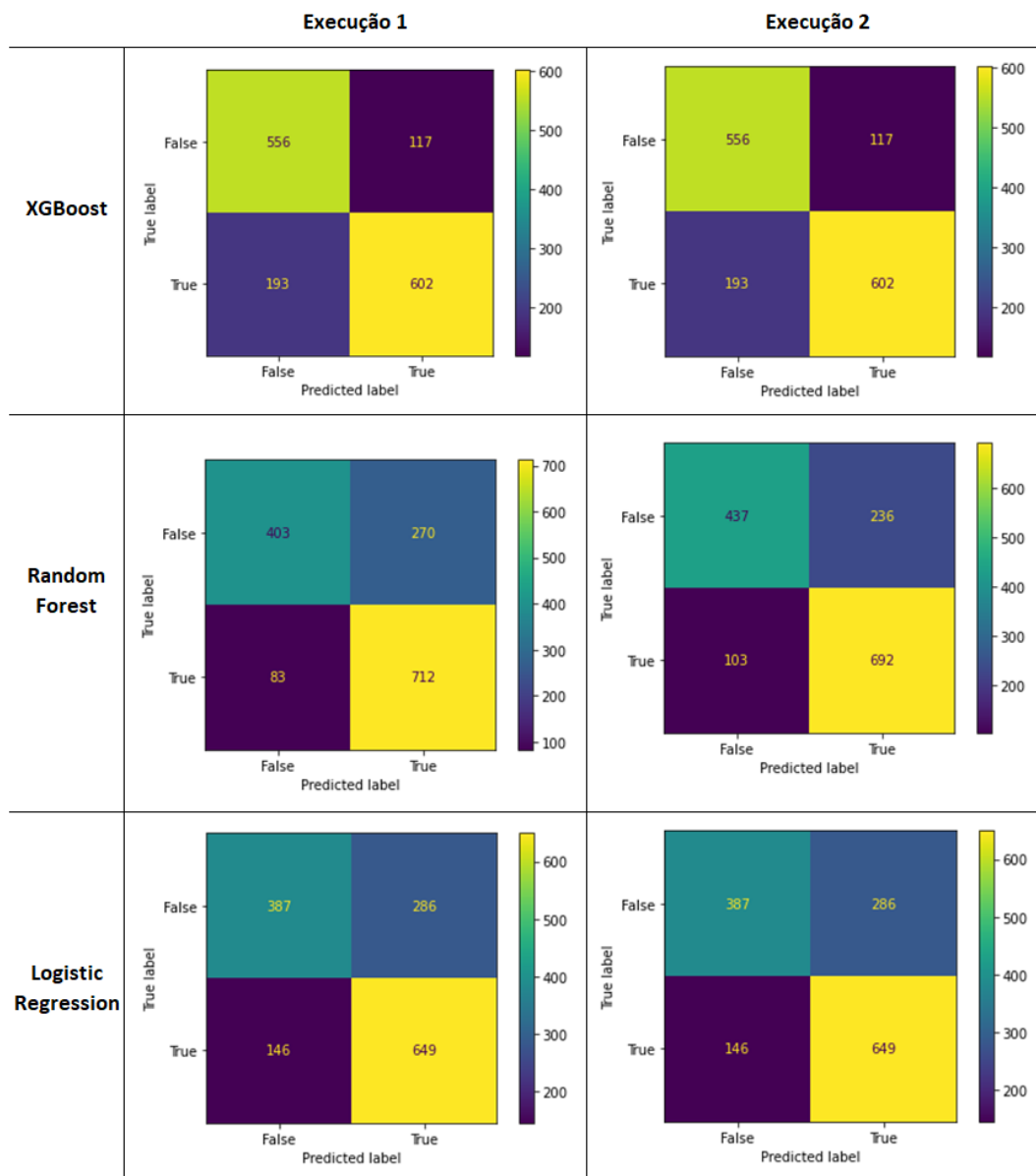


Figura 15 - Matrizes de Confusão do Experimento 1.

### 5.5 Experimentos 2 e 3

Os Experimentos 2 e 3, conforme a Tabela 4, são do segundo bloco de objetivos intermediários e buscam verificar como informações sobre o histórico das decisões do respondente pode interatuar com o modelo de predição proposto. Para tanto, foi criada na etapa de *feature engineering* uma nova coluna binária com a informação de se a última decisão tomada pelo decisor foi com viés do tipo *Risk-Seeking*, denominada *RS\_before*.

Mais especificamente, o Experimento 2 irá adicionar o atributo de ocorrência de viés cognitivo de busca de risco na decisão anterior, para avaliar se a acurácia será modificada com essa nova variável. De maneira semelhante, o

Experimento 3 irá acrescentar a variável de Feedback da decisão, ou seja, as decisões que ao final eram apresentados os resultados aos respondentes.

Sendo assim, as variáveis usadas em cada um dos Experimentos 2 e 3 estão elencadas na Tabela 12, os novos cenários comparados neste experimento foram:

- Cenário 3: com a variável de *Risk Seeking* da decisão anterior (RS\_Before);
- Cenário 4: com as variáveis *Risk Seeking* da decisão anterior (RS\_Before) e a de *Feedback* recebido pelos respondentes.

O sumário das variáveis usadas e as melhores configurações de parâmetros para cada modelo testado encontram-se na Tabela 13. Em uma análise dos dados, é possível verificar que 39% dos respondentes escolheram uma opção com viés. No entanto, quando um respondente possui assinalado o atributo RS\_before, ou seja, escolheu a opção enviesada na última decisão, o percentual de escolha do viés cognitivo é de 77%, indicando a manutenção de uma posição propensa a risco mesmo que isso já tenha ocorrido na decisão anterior.

Cenários / Variáveis	HighValue	pHighValue	LowValue	pLowValue	RS_before	Feedback
Cenário 3	X	X	X	X	X	
Cenário 4	X	X	X	X	X	X

Tabela 12 - Variáveis explicativas dos Experimentos 2 e 3.

A Busca Randomizada Bayesiana também foi aplicada nesse cenário como estratégia de varredura de parâmetros, com os melhores resultados elencados na Tabela 13. Observa-se que foram otimizados para valores diferentes do que o primeiro Experimento, demonstrando que para convergir a um resultado ótimo foram necessárias mudanças nesses atributos.

Modelos	Melhores Parâmetros	
	Cenário 3	Cenário 4
Random Forest	criterion: gini, max_depth: 50, n_estimators: 53, max_features: 86, min_samples_leaf: 72, min_samples_split: 4	criterion: gini, max_depth: 50, n_estimators: 66, max_features: 78, min_samples_leaf: 5, min_samples_split: 2
XGBoost	learning_rate: 0.01119, max_depth: 30, n_estimators: 30, subsample: 0.55769, colsample: 0.92257, min_child_weight: 89	learning_rate: 0.00163, max_depth: 98, n_estimators: 22, subsample: 0.78740, colsample: 0.90115, min_child_weight: 94
Logistic Regression	C: 0.0001, fit_intercept: False, penalty: l2, solver: lbfgs	C: 0.0001, fit_intercept: True, penalty: l2, solver: liblinear

Tabela 13 - Melhores hiperparâmetros por Modelo dos Experimento 2 e 3.

Em relação aos modelos de predição a variável *RS\_before* não impactou a acurácia geral dos modelos de predição, demonstrado na Figura 16. Como é possível observar no Resumo dos resultados, no Cenário 3, a única taxa que foi sensibilizada pela nova variável foi a do Random Forest, subindo para 76,91% com relação ao modelo base com apenas os valores e as probabilidades, o mesmo percentual alcançado pelo Cenário 2.

De maneira semelhante ao Experimento 1, também foram calculadas as acurácias por classes de ganhos e perdas, conforme a Tabela 14. Através dela pode-se constatar também a mesma distribuição observada nas execuções e cenários anteriores com o XGBoost predizendo melhor as perdas e o *Random Forest* os ganhos.

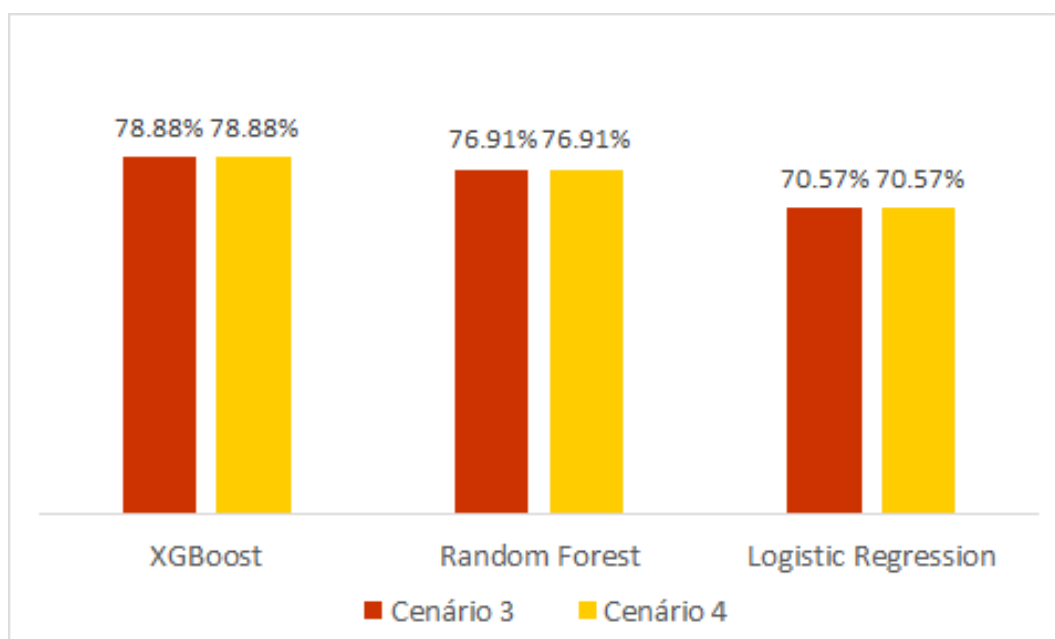


Figura 16 - Acurácia dos Modelos por Cenário no Experimento 2 e 3.



Como ocorreu no Experimento 1, ao olhar para a acurácia dos modelos por classe da variável alvo, nota-se a mesma distribuição ocorrida no experimento anterior, disponíveis na Tabela 14 de acurácia por classes. Os melhores resultados de predição para a classe de ganhos sendo do XGBoost enquanto a classe de perdas foi mais bem predita pelo Random Forest, com 83,73% e 80,93% respectivamente. A Regressão Logística obteve a acurácia geral de 70,57%, seguindo o padrão do Experimento 1.

Para complementar essa análise, foram observados o comportamento de escolha dos decisores após receberam um *feedback* de sua escolha, isto é, puderam ver o resultado que obtiveram e o resultado da opção contrária. Dos *Trials* de 1 a 5, onde não há o conhecimento da resposta do experimento pelo decisor, o percentual de escolha da opção enviesada foi de aproximadamente 42%. Nos *Trials* seguintes, cientes das respostas de suas escolhas, essa taxa caiu 2 pontos percentuais.

Embora esse percentual de mudança nas respostas não tenha sido muito alto, o aproveitamento das decisões aumentou ainda mais. A taxa de ganhos antes do *Feedback* era de 55% e cresceu para 60% após a informação dos decisores, indicando que na média obtiveram melhores retornos depois de conhecerem as respostas ocorridas nos *Games*.

Então, o *feedback* foi incorporado à base de dados, juntamente com o indicador de viés na decisão anterior, para verificar se haveria mudança na predição do modelo de ganhos e perdas no Cenário 4. Os melhores resultados atingidos pelos modelos seguiram os mesmos, tanto para a acurácia geral como para a acurácia por classe de ganho e perda nos três modelos testados.

Acurácia					Matriz de Confusão				
Cenários		Classe=Perda	Classe=Ganho	Total	TP	FP	TN	FN	Total
Cenário 3	XGBOOST	74.23%	83.73%	78.88%	602	117	556	193	1468
	Random Forest	80.93%	74.57%	76.91%	692	236	437	103	1468
	Regressão Logística	72.61%	69.41%	70.57%	649	286	387	146	1468
Cenário 4	XGBOOST	74.23%	83.73%	78.88%	602	117	556	193	1468
	Random Forest	80.93%	74.57%	76.91%	692	236	437	103	1468
	Regressão Logística	72.61%	69.41%	70.57%	649	286	387	146	1468

Tabela 14 - Acurácia por classe nos Experimentos 2 e 3.

Todas as matrizes de confusão geradas para os Experimentos 2 e 3 estão na Figura 17, trazendo graficamente a concentração em cada categoria de

predição das classes atribuídas. As cores mais claras significam concentrações maiores, enquanto as escuras concentrações menores.

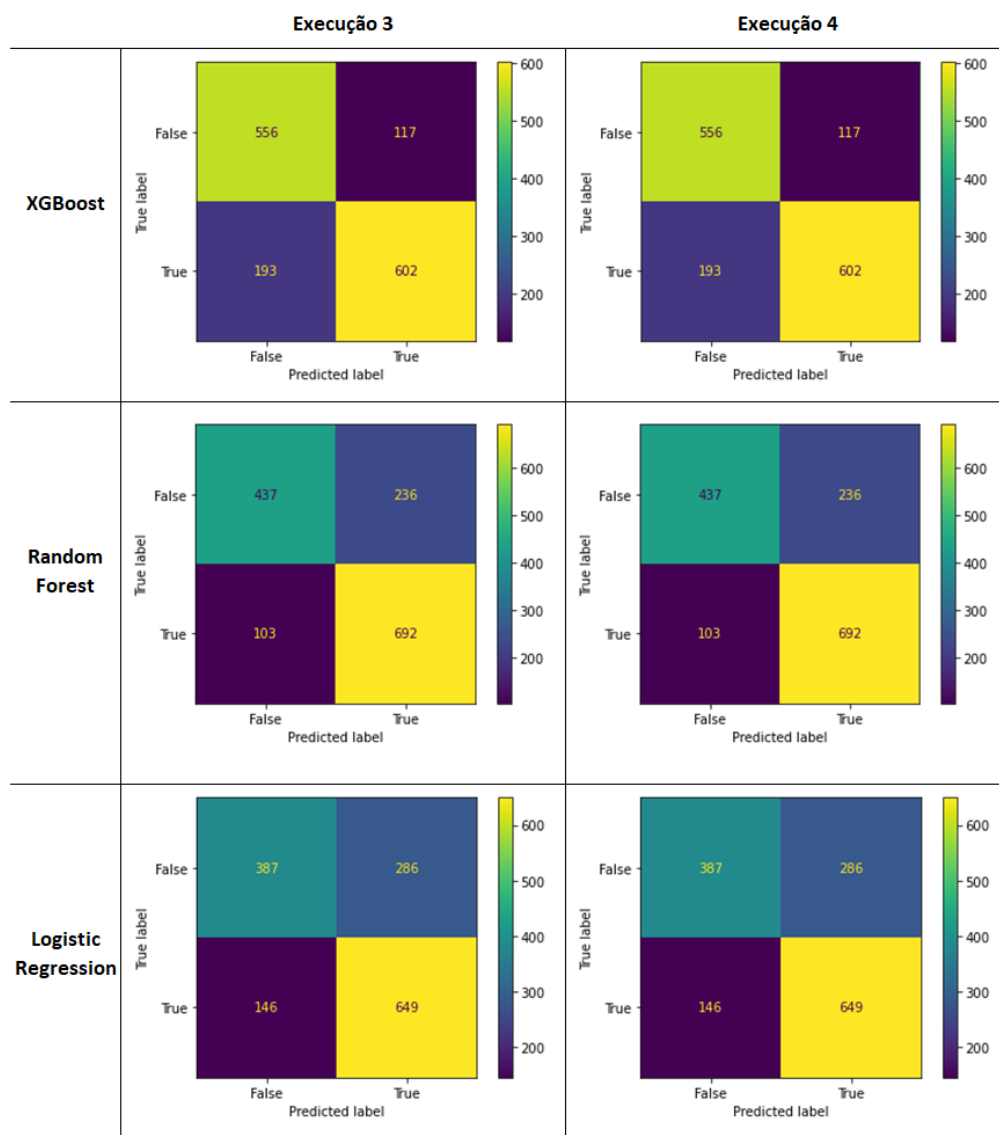


Figura 17 - Matrizes de Confusão dos Experimentos 2 e 3.

## 6. Interpretações e Discussões

Um fator interessante ao se analisar nos dados de avaliação do Experimento 1 é como a acurácia total do Random Forest teve um incremento, apesar de sua classe de melhor predição – as perdas –, terem uma redução. Essa percepção se deu devido à classe de ganhos ter um incremento percentual ligeiramente maior e porque essa classe representa uma proporção um pouco maior dos dados, como mostrado anteriormente. Sendo assim, esse maior equilíbrio das acurácias de ganhos e perdas do Random Forest fez com que sua acurácia geral tivesse um incremento de aproximadamente 1 ponto percentual.

Observa-se, também, que, apesar da Regressão Logística ter tido mais acertos da classe de ganhos do que o XGBoost em valores absolutos, a diferença percentual foi de quase 15 pontos percentuais a favor do modelo com menos acertos. Isso ocorre pois a Regressão também errou mais do que o dobro de casos do XGBoost para classificação dos ganhos, o que percentualmente fez essa grande diferença.

No cenário 2, a diferença observada foi a mudança apenas para o Random Forest, que nos casos de ganho registrou um percentual de acerto de 80,93% e nos de perda 74,57%. Os percentuais ficaram mais próximos entre si do que no primeiro cenário. Ainda assim, sua acurácia geral subiu quase um ponto percentual, indicando uma possível melhora no modelo.

Com relação às opções escolhidas pelos respondentes do CPC, cerca de 39% das respostas foram com o viés. Dentre os respondentes que selecionaram uma opção enviesada, aproximadamente 48% obtiveram ganhos, enquanto os 52% restantes tiveram perdas. Isso demonstra que, apesar de terem selecionados opções mais arriscadas, os decisores tiveram quase o mesmo aproveitamento relativo em termos de número de vezes em que ganharam e perderam. Apesar de uma decisão enviesada ser prejudicial, ela ainda depende do fator da probabilidade, que pode influenciar nessa estatística geral.

Para os Experimentos 2 e 3, as características históricas do decisor, *features RS\_before* e *feedback*, não foram determinantes na piora ou melhora do modelo de predição, com relação às suas acurácias. A partir da amostra desses resultados, infere-se que os modelos conseguiram identificar sem a nova *feature*

quando haveria influência do viés cognitivo de busca de risco. Por isso, esses atributos não foram determinantes para a predição dos ganhos ou das perdas.

Outros vieses, portanto, também podem estar relacionados ao aumento da taxa de ganhos após a implementação do *feedback*, como o Viés de Excesso de Confiança, o Efeito de Recência e o Viés Retrospectivo. Em todos esses três vieses, o entendimento do que ocorreu em decisões passadas é interpretado de maneira errônea/enviesada quando se depara com novas escolhas a serem feitas.

Apesar de cada *game* ser independente do outro no que diz respeito às suas probabilidades, quando se tem uma experiência prévia essas percepções podem ser distorcidas para gerar uma sensação de segurança e conformidade ao escolher determinada opção, mesmo que não seja baseada em características reais do problema.

## 7. Conclusão

Apesar da sua crescente importância e impacto na tomada de decisões do negócio, há poucos trabalhos analisando o impacto de vieses cognitivos na literatura, e que geralmente endereçam apenas o problema específico para o qual foram estabelecidas, ainda que rastreiem vários tipos de vieses cognitivos. Importante observar ainda que, quanto maior a generalização e abrangência, mais difícil torna-se a detecção de muitos vieses com qualidade e fluidez.

Portanto, o objetivo desta pesquisa foi propor uma abordagem orientada a dados para analisar os vieses cognitivos presentes nas decisões, de maneira a dar maior autonomia e ciência ao decisor sobre suas preferências de risco e vieses aos quais está sujeito. Para isso, foi proposta uma arquitetura que serve como base para um sistema de apoio a decisão onde todas as informações são armazenadas e fornecidas ao decisor em tempo real e em forma de análise após um histórico de decisões.

A arquitetura proposta nesta pesquisa estabelece como todas as etapas envolvidas (desde a concepção de um possível problema a ser abordado até as inferências baseadas nos dados) podem interagir e se complementar, dando maior robustez às análises e direcionando os possíveis desdobramentos.

A arquitetura proposta, e em particular o seu módulo de Aprendizagem do Impacto dos Vieses Cognitivos pelos Dados, foram avaliados em 3 experimentos em cenários envolvendo perdas monetárias, que visaram a criação de um modelo preditivo de ganhos e perdas utilizando vieses cognitivos de busca de risco como fonte de aprimoramento e interpretabilidade do modelo.

Os resultados da experimentação permitiram concluir que o simples atributo de viés cognitivo na base não influenciou na performance do modelo de predição. Isso pode ser atribuído ao fato de que para calcular se uma opção é enviesada ou não são utilizados os mesmos dados da base. Quando são adicionados atributos históricos na base, de viés na decisão anterior e de *feedback* recebido pelo decisor, apesar dos usuários terem aumentado seus percentuais de ganho, pouco fez diferença na acurácia dos modelos de predição. Observou-se, ainda, que determinados modelos são mais especializados para prever as perdas e outros a prever os ganhos, segundo as acurácias por classe determinadas.

Este trabalho pode ser replicado para investidores individuais ou organizações, aproveitando o arcabouço teórico e o encadeamento lógico da atribuição de valor de ganho ou perda às decisões, bem como a parte técnica de implementação dos modelos e modelagem dos dados.

As organizações podem utilizar a arquitetura para explorar os vieses cognitivos de busca de risco presentes nos processos de negócios, utilizando os próprios dados para auxiliar os tomadores de decisão. Assim, podem maximizar ganhos ou diminuir suas perdas por más escolhas feitas, tornando o processo mais ciente dos riscos e vieses presentes.

As contribuições desta pesquisa incluem:

- Do ponto de vista científico, a caracterização de todo o cenário baseado em Ciência de Dados e Economia Comportamental para tornar o processo de tomada de decisão mais consciente de seus vieses, no formato da arquitetura proposta. Essa base pode servir como um guia para um sistema de apoio à decisão computacional e com os recursos e funções especificados na arquitetura e no módulo de aprendizagem dos padrões.

- Do ponto de vista tecnológico, as contribuições dos modelos preditivos corroboram a aptidão dos modelos para as técnicas de classificação e a otimização dos hiperparâmetros com os algoritmos foi disponibilizada no Github da pesquisa para difundir ainda mais esse conhecimento.

Os resultados desta pesquisa possuem algumas limitações, a serem descritas a seguir, como: o cenário de aplicação, os modelos utilizados, as variáveis disponíveis e o viés tratado. O cenário de avaliação do Módulo de Aprendizagem aplicado em apenas um cenário, em virtude dos dados disponíveis para análise. Recomenda-se avaliar aplicação em outros cenários que possam oferecer os dados e o seu respectivo contexto.

Os modelos de aprendizado de máquina que foram testados são apenas três, pode-se incluir outros para testes semelhantes, como por exemplo redes neurais. Os três modelos escolhidos servem como benchmark por serem três dos modelos mais conhecidos para as técnicas de classificação de aprendizado supervisionado.

O tipo de viés considerado na aplicação possui um formato específico e em virtude de o contexto da área de estudo apresentado precisa de características que podem limitar determinadas aplicações. No caso de outros tipos de vieses cognitivos forem abordados, recomenda-se que sejam desenvolvidas outras ontologias para explicitar semanticamente os conceitos aplicados. Então, dentre

diversos benefícios, será possível automatizar e documentar os processos de classificação de vieses nos casos de estudo.

Trabalhos futuros podem incluir a aplicação de simulação baseada em agentes (Kazil, et al., 2020) para estimar quais os parâmetros mais sensíveis para a detecção dos vieses cognitivos em diversos problemas. Outras aplicações possíveis são de técnicas de aprendizagem por reforço e por reforço inverso, onde é possível que o modelo aprenda a função de utilidade presente em sua amostra de treino, ou ainda quais as ações necessárias para garantir o atendimento das necessidades do decisor (Peysakhovich, 2019).

Desdobramentos futuros incluem variáveis alvo que não sejam binárias, para quantificar o tamanho dos ganhos e das perdas, juntamente com outros tipos de vieses cognitivos que podem ser explorados. A partir do momento que houver uma lógica que possa ser transmitida para uma linguagem matemática, como uma ontologia, mesmo que seja particular do domínio de aplicação, pode-se tratar outros vieses cognitivos mais abstratos e pouco explorados na literatura.

## 8. Referências Bibliográficas

Arnott, D. (2006). Cognitive biases and decision support systems development: a design science approach. *Information Systems Journal*, 16(1), 55-78.

Arnott, D., & Gao, S. (2019). Behavioral economics for decision support systems researchers. *Decision Support Systems*, 122, 113063.

Barber, J. P. (2009). Toward a working through of some core conflicts in psychotherapy research. *Psychotherapy research*, 19(1), 1-12.

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of machine learning research*, 13(2).

Bourgin, D. P., Peterson, J.C., Reichman, D., Russel, S.J., Griffiths, T. (2019). Cognitive Model Priors for Predicting Human Decisions. Conference: International Conference on Machine Learning

Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32.

Broihanne, M. H., Merli, M., & Roger, P. (2014). Overconfidence, risk perception and the risk-taking behavior of finance professionals. *Finance Research Letters*, 11(2), 64-73.

Chen, T. et al. Xgboost: extreme gradient boosting. R package version 0.4-2, v. 1, n. 4, 2015.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., ... & Zhou, T. (2015). Xgboost: extreme gradient boosting. R package version 0.4-2, 1(4), 1-4.

Cohen M., Etner, J., & Jeleva, M. (2008). Dynamic decision making when risk perception depends on past experience. *Theory and Decision*, 64(2–3), 173–192. <https://doi.org/10.1007/s11238-007-9061-3>



Edwards, W. (1954). The theory of decision making. *Psychological bulletin*, 51(4), 380.

Engelmann, J. B., Damaraju, E., Padmala, S., & Pessoa, L. (2009). Combined effects of attention and motivation on visual task performance: transient and sustained motivational effects. *Frontiers in human neuroscience*, 3, 342.

Fischhoff, B. (1975). Hindsight not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 288–299. <https://doi.org/10.1136/qhc.12.4.304>

Frawley, W. J., Piatetsky-Shapiro, G., & Matheus, C. J. (1992). Knowledge discovery in databases: An overview. *AI magazine*, 13(3), 57-57.

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing?. *International journal of human-computer studies*, 43(5-6), 907-928.

Guizzardi, G., & Halpin, T. (2008). Ontological foundations for conceptual modelling. *Applied Ontology*, 3(1-2), 1-12.

Guizzardi, R., Carneiro, B. G., Porello, D., & Guizzardi, G. (2020, November). A core ontology on decision making. In *Proceedings of the XIII Seminar on Ontology Research in Brazil and IV Doctoral and Masters Consortium on Ontologies (ONTOBRAS 2020)* (Vol. 2728, pp. 9-21). CEUR-WS.

Haselton, M. G., Nettle, D., & Andrews, P. W. (2015). The evolution of cognitive bias. *The handbook of evolutionary psychology*, 724-746.

Hassan, N. Y., Gomaa, W. H., Khoriba, G. A., & Haggag, M. H. (2020). Credibility Detection in Twitter Using Word N-gram Analysis and Supervised Machine Learning Techniques. *International Journal of Intelligent Engineering & Systems*, 13(1).

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. 398). John Wiley & Sons.

Huyn, P., Kumar, A., Peng, S., Personalizing Consumer Interventions Through Leveraging Behavioral Economics, *Procedia Computer Science*, Volume 207, 2022, Pages 282-291, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2022.09.061>

Kahneman, D. (2003). Maps of bounded rationality: Psychology of behavioral economics. *The American Economic Review*, 93(5), 1449–1475.

Kahneman, D., & Tversky A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–292.

Kahneman, D., & Tversky A. (1984). Choices, values, and frames. *American Psychologist*. <https://doi.org/10.1037/0003-066X.39.4.341>

Kahneman, D., Knetsch J. L., & Thaler R. H. (1989). Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy*, 79(5), 1277–1284.

Kahneman, D., Knetsch J. L., Thaler R. H., Johnson H., & Professor L. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *The Journal of Economic Perspectives*, 5(1), 193–206.

Kahneman, D., *Thinking, fast and slow*. Macmillan. 2011

Kumari, R., & Srivastava, S. K. (2017). Machine learning: A review on binary classification. *International Journal of Computer Applications*, 160(7).

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1992). The impact of institutional trading on stock prices. *Journal of financial economics*, 32(1), 23-43.

Levine, H. A. (2007). *Project Portfolio Management: A Practical Guide to Selecting Projects, Managing Portfolios, and Maximizing Benefits*. San Francisco, CA.: John Wiley & Sons

Lewandowski, M. Prospect theory versus expected utility theory: Assumptions, predictions, intuition and modelling of risk attitudes. *Central European Journal of Economic Modelling and Econometrics*, 2017, 275-321-275-321.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news*, 2(3), 18-22.

Lindblom, C. E. (1979). Still muddling, not yet through. *Public administration review*, 39(6), 517-526.

Lord C. G., Ross L., & Lepper M. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37(11), 2098–2109.

Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*. [Internet], 9(1), 381-386.

March, J. G. (1962). The business firm as a political coalition. *The Journal of politics*, 24(4), 662-678.

March, J. G. (1978). Bounded rationality, ambiguity, and the engineering of choice. *The bell journal of economics*, 587-608.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR)*, 54(6), 1-35.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM computing surveys (CSUR)*, 54(6), 1-35.

Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H., & Kassam, K. S. (2015). Debiasing Decisions: Improved Decision Making With a Single Training Intervention. *Policy Insights from the Behavioral and Brain Sciences*, 2(1), 129–140.

Nelli, F., & Nelli, F. (2015). Machine Learning with scikit-learn. Python Data Analytics: Data Analysis and Science Using Pandas, matplotlib, and the Python Programming Language, 237-264.

Netto, A., & Maciel, F. (2021). Python Para Data Science: E Machine Learning Descomplicado. Alta Books.

Nobre, F.C., Machado, M.J.C. and Nobre, L.H.N. Behavioral Biases and the Decision-Making in Entrepreneurs and Managers. Revista de Administração Contemporânea [online]. 2022, v. 26, suppl 1

Noy, N. F., & McGuinness, D. L. (2001). Ontology development 101: A guide to creating your first ontology.

Paiva, F. D., Cardoso, R. T. N., Hanaoka, G. P., & Duarte, W. M. (2019). Decision-making for financial trading: A fusion approach of machine learning and portfolio selection. Expert Systems with Applications, 115, 635-655.

Paulus, D., Fathi, R., Fiedrich, F., de Walle, B. V., & Comes, T. (2022). On the Interplay of Data and Cognitive Bias in Crisis Information Management. Information Systems Frontiers, 1-25.

Pfeffer, J. (1981). Understanding the role of power in decision making. Power in organizations, 404, 423.

Power, D. J., & Kaparthy, S. (2002). Building Web-based decision support systems. Studies in Informatics and Control, 11(4), 291-302.

Priebe, T., & Markus, S. (2015, October). Business information modeling: A methodology for data-intensive projects, data science and big data governance. In 2015 IEEE International Conference on Big Data (Big Data) (pp. 2056-2065). IEEE.

Provost, F., & Fawcett, T. (2013). Data Science for Business: What you need to know about data mining and data-analytic thinking. " O'Reilly Media, Inc."

Raschka, S., & Mirjalili, V. (2019). Learning best practices for model evaluation and hyperparameter tuning. Python Mach. Learn., Third, Packt Publishing, 280-321.

Roese, N. J., & Olson, J. M. (2007). Better, stronger, faster: Self-serving judgment, affect regulation, and the optimal vigilance hypothesis. Perspectives on Psychological Science, 2(2), 124-141.

Roese, N. J., & Vohs, K. D. (2012). Hindsight bias. Perspectives on psychological science, 7(5), 411-426.

Samuelson W., & Zeckhauser R. (1988). Status quo bias in decision making. Journal of Risk and Uncertainty, 1(1), 7–59.

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. Journal of risk and uncertainty, 1, 7-59.

Scharfstein, D. S., & Stein, J. C. (1990). Herd behavior and investment. The American economic review, 465-479.

Sellier, A. L., Scopelliti, I., & Morewedge, C. K. (2019). Debiasing training improves decision making in the field. Psychological science, 30(9), 1371-1379.

Sherif M., Taub D., & Hovland C. I. (1958). Assimilation and contrast effects of anchoring stimulating on judgments. Journal of Experimental Psychology, 55(2), 150–155.

Simon, H. A. (1957). Models of man; social and rational.

Simon, H. A. (1977). The logic of heuristic decision making. In Models of Discovery: and Other Topics in the Methods of Science (pp. 154-175). Dordrecht: Springer Netherlands.

Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. Advances in neural information processing systems, 25.

Studer, R., Benjamins, V. R., & Fensel, D. (1998). Knowledge engineering: Principles and methods. *Data & knowledge engineering*, 25(1-2), 161-197.

Thaler R. H. (1985). Mental accounting and consumer choice. *Marketing Science*, 4(3), 199–214.

Thaler R. H., & Johnson E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Science*. <https://doi.org/10.1287/mnsc.36.6.643>

Thaler, R.H. and Ganser, L.J., *Misbehaving: The making of behavioral economics*. New York: WW Norton. 2015.

Turpin, S. M., & Marais, M. A. (2004). Decision-making: Theory and practice. *ORiON*, 20(2), 143-160.

Tversky A., & Kahneman D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, 76(2), 105–110.

Tversky A., & Kahneman D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106(4), 1039–1061.

Tversky A., & Kahneman D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*. <https://doi.org/10.1007/BF00122574>

## 9. Apêndice

### Dicionário para nomes de variáveis em dados brutos CPC1

**SubjID** – Identificador único do sujeito humano, composto por 5 dígitos. O 1º dígito marca o conjunto de jogos que o sujeito enfrentou (1-7), o 2º dígito marca o local em que o sujeito jogou (0-1), os últimos três dígitos não têm significado.

**Location** – A localização física do sujeito participante (“Technion”/“Rehovot”).

**Sex** - sexo do sujeito (M/F).

**Age** - idade do sujeito no momento do experimento.

**Set** – ID do número do set que o sujeito enfrentou (cada set consiste nos mesmos 30 jogos). 1-5 são dados do CPC15, 6-7 são dados do Experimento 1 do CPC18.

**Condition** – Variável herdada do CPC15 (“ByProb”/“ByFB”). Refere-se à condição do pedido por qual sujeito viu os jogos dentro de um conjunto. Os dados do CPC18 estarão todos na condição “ByProb”.

**GameID** – Identificador exclusivo do jogo (problema de escolha) (1-210).

<< As próximas 12 variáveis definem (conjuntamente) o problema de escolha, definindo as duas possíveis distribuições e qualquer relação entre elas. Para obter mais detalhes sobre como cada 12 tuplas define as distribuições exatas, veja o papel >>

**Ha** – Valor esperado da loteria (Alta) na Opção A.

**pHa** – Probabilidade de obter o pagamento (Alto) da loteria na Opção A.

**La** – Retorno baixo na Opção A.

**LotShapeA** – Forma da loteria na Opção A (“-“, “Symm”, “L-skew” ou “R-skew”).

**LotNumA** - Número de resultados da loteria na Opção A.

**Hb** – Valor esperado da loteria (Alta) na Opção B.

**pHb** - Probabilidade de obter o pagamento (Alto) da loteria na Opção B.

**Lb** - Baixo retorno na Opção B.

**LotShapeB** – Forma da loteria na Opção B (“-“, “Symm”, “L-skew” ou “R-skew”).

**LotNumB** - Número de resultados da loteria na Opção B.

**Amb** – Se a Opção B é ambígua (ou seja, suas probabilidades não são descritas aos sujeitos; Boleano).

Corr – Se os payoffs gerados pelas duas opções possíveis estão correlacionados e o sinal da correlação (-1/0/1).

*Order* – A posição serial do jogo atual dentro da sequência de 30 jogos que o sujeito enfrentou (1-30).

*Trial* – O número da tentativa dentro de um jogo (1-25).

Button – O lado na tela do botão escolhido (“L”/”R”).

B – A variável de resposta. Se o sujeito selecionou ou não a Opção B no ensaio atual (Booleano).

*Payoff* – O *payoff* que o sujeito obteve de sua escolha no teste atual.

*Forgone* – A recompensa que o sujeito teria obtido se ele/ela tivesse selecionado a outra opção no julgamento atual.

RT – Tempo de reação até a escolha da opção (em milissegundos). Medido apenas para as séries 6, 7.

<< As próximas quatro variáveis podem ser calculadas diretamente das variáveis anteriores. eles são dados Por conveniência. >>

APay - O pagamento fornecido pela Opção A no teste atual.

BPay - O pagamento fornecido pela Opção B no teste atual.

*Feedback* – Se o *feedback* (completo) foi fornecido para o sujeito em relação aos pagamentos no atual julgamento.

*Block* – Número de blocos de tempo dentro do jogo atual (cada 5 tentativas definem um bloco).