

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO

**Análise de Satisfação de Serviços e Produtos
Brasileiros em Plataformas Eletrônicas: Uma
Abordagem de Mineração de Texto**

Isabella Lopes Carmo dos Santos

PROJETO FINAL DE GRADUAÇÃO

CENTRO TÉCNICO CIENTÍFICO - CTC

DEPARTAMENTO DE INFORMÁTICA

Curso de Graduação em Ciência da Computação

Rio de Janeiro, Junho de 2024



Isabella Lopes Carmo dos Santos

**Análise de Satisfação de Serviços e Produtos
Brasileiros em Plataformas Eletrônicas: Uma
Abordagem de Mineração de Texto**

Relatório de Projeto Final, apresentado ao programa Ciência da Computação da PUC-Rio como requisito parcial para a obtenção do título de Bacharel em Ciência da Computação.

Orientador: Marcos Vianna Villas

Departamento de Informática

Rio de Janeiro

Junho de 2024

AGRADECIMENTOS

Agradeço a Deus e a Nossa Senhora por todas as bênçãos recebidas ao longo do meu período na universidade, guiando sempre o meu caminho.

Aos meus amados pais, expresso minha mais profunda gratidão por me proporcionarem a oportunidade de estudar na PUC-Rio e por acreditarem incondicionalmente em mim. Vocês são o meu maior presente, a motivação que me move e o porto seguro que me acolhe em qualquer maré. Amo-os imensamente.

Ao meu estimado orientador, Marcos Vianna Villas, registro meus sinceros agradecimentos. Desde o início apoiou a minha ideia e me auxiliou da melhor forma possível, sendo essencial para a construção deste trabalho.

Minha profunda gratidão estende-se a todos os meus professores que me guiaram desde o início da minha jornada acadêmica. Muitos se tornaram modelos para minha vida profissional e acadêmica, especialmente os professores Marcos Villas e Sérgio Lifschitz, que acreditaram em meu potencial desde o primeiro momento em que estagiei no Laboratório BioBD, proporcionando-me um ambiente de aprendizado e crescimento inigualável.

Aos meus queridos amigos de infância, Esther Obriem, Leonardo Santana, Luiz Henrique Sobral, Luíza Catunda e Rebecca Chiavazzoli, agradeço por celebrarem todas as minhas conquistas e por permanecerem ao meu lado, oferecendo-me força e apoio incondicional em inúmeros momentos, muitas vezes sem sequer perceberem. A amizade de vocês é uma dádiva e sou imensamente grata por todo carinho.

Aos amigos que a faculdade me presenteou, amizades que levarei em meu coração para sempre, expresso meu profundo agradecimento. Vocês me acompanharam de perto, tornando cada momento na faculdade especial e repleto de alegria. Ao meu namorado, minha inspiração constante e um dos meus maiores apoiadores, André Luiz, aos meus amigos Alexandre Augusto, Bernardo Delgado, Gustavo Sampaio, Laura Luz, Luca Ribeiro, Luiza Bretas, agradeço a cada um de vocês por cada gesto de companheirismo e incentivo.

Serei eternamente grata pelos estudos na PUC-Rio e por construir memórias incríveis com todos citados acima. Que venham os próximos desafios.

RESUMO

Lopes Carmo dos Santos, Isabella. Vianna Villas, Marcos. **Análise de Satisfação de Serviços Brasileiros em Plataformas Eletrônicas**. Rio de Janeiro, 2024. 126 p. Relatório de Projeto Final – Departamento de Informática. Pontifícia Universidade Católica do Rio de Janeiro.

Esse projeto de graduação visa desenvolver uma metodologia para avaliação da satisfação dos clientes presentes em plataformas eletrônicas relacionadas a serviços e produtos brasileiros, utilizando técnicas de mineração de texto. O método identificará tópicos, sentimentos e entidades relevantes presentes nos comentários dos consumidores. A solução proposta tem a capacidade de auxiliar empresas e comerciantes a aprimorar seus produtos, serviços e atendimento ao cliente.

PALAVRAS-CHAVE

Satisfação do cliente; Mineração de texto; Plataformas eletrônicas; Metodologia.

ABSTRACT

Lopes Carmo dos Santos, Isabella. Vianna Villas, Marcos. **Satisfaction Analysis of Brazilian Services on Electronic Platforms**. Rio de Janeiro, 2024. 126 p.
Relatório de Projeto Final – Departamento de Informática. Pontifícia Universidade Católica do Rio de Janeiro.

This graduation project aims to develop a methodology for evaluating customer satisfaction presented on electronic platforms related to Brazilian services and products, using text mining techniques. The method will identify relevant topics, feelings and entities present in consumer comments. The proposed solution has the ability to help companies and merchants to improve their products, services and customer service.

KEYWORDS

Customer satisfaction; Text mining; Electronic platforms; Methodology.

SUMÁRIO

1.	INTRODUÇÃO.....	7
2.	SITUAÇÃO ATUAL	7
3.	OBJETIVO E PROPOSTA DO TRABALHO	9
4.	PLANO DE AÇÃO	10
5.	LEVANTAMENTO DOS MÉTODOS ATUAIS	11
6.	ESTUDO DE MINERAÇÃO DE TEXTO	14
6.1.	PRÉ-PROCESSAMENTO.....	15
6.2.	MINERAÇÃO DE TEXTO	16
6.3.	AValiação CONCLUSIVA	18
7.	SELEÇÃO DE TÉCNICAS PARA APLICAÇÃO NOS CASOS.....	18
8.	SELEÇÃO E DESCRIÇÃO DOS CASOS	20
9.	CONSTRUÇÃO DO MÉTODO	26
10.	APLICAÇÃO DAS TÉCNICAS ESCOLHIDAS	31
10.1.	RESTAURANTE GUACAMOLE.....	32
10.2.	OLIST.....	48
10.3.	BUSCAPÉ.....	58
10.4.	B2W DIGITAL	74
10.5.	HOTÉIS BRASILEIROS.....	86
10.6.	HOTEL GRAND HYATT DE SÃO PAULO	102
10.7.	SMARTPHONE DA MOTOROLA	120
11.	AValiação TÉCNICA DO MÉTODO E SEUS RESULTADOS	126
12.	AValiação DA UTILIDADE DO MÉTODO POR ESPECIALISTAS	129
13.	CONSIDERAÇÕES FINAIS.....	132
14.	REFERÊNCIAS BIBLIOGRÁFICAS.....	135
15.	ANEXOS	139

1. INTRODUÇÃO

O crescimento exponencial do comércio eletrônico no Brasil tem gerado um ambiente altamente competitivo, onde a satisfação do cliente desempenha um papel crítico no sucesso das empresas. A compreensão das opiniões e sentimentos dos consumidores após a prestação de serviços e vendas de produtos de tais empresas é essencial para melhorar a qualidade de seus serviços, produtos e atendimento ao cliente. É evidente a necessidade que as empresas tenham um insight sobre toda essa questão.

Este projeto tem como objetivo desenvolver uma metodologia e escolher algoritmos para avaliar a satisfação dos clientes em relação a serviços e produtos oferecidos no contexto digital brasileiro, empregando técnicas de mineração de texto. O desenvolvimento do mesmo incluirá o levantamento de tais técnicas e a coleta de dados (comentários após a prestação de serviços e aquisição de produtos) de serviços e produtos brasileiros em plataformas eletrônicas para desenvolver uma solução capaz de identificar os sentimentos dos clientes (satisfação ou insatisfação) em relação a tais procedimentos. Além disso, serão identificados os tópicos relevantes de discussão e tendências emergentes por meio de algoritmos e a identificação de entidades presentes nos documentos. A metodologia desenvolvida e os algoritmos escolhidos poderão ser utilizados por empresas que buscam melhorar seus produtos, serviços e estratégias de atendimento ao cliente.

2. SITUAÇÃO ATUAL

A expansão significativa do comércio eletrônico brasileiro (Mundo do Marketing, 2023) tem gerado um ambiente altamente competitivo, onde a satisfação do cliente desempenha um papel crítico no sucesso das empresas no mercado virtual. A satisfação do cliente reflete a medida em que suas expectativas são atendidas ou superadas em relação aos produtos, serviços e experiência de compra online.

Em um ambiente competitivo de prestação de serviços e produtos no cenário digital, oferecer um serviço de qualidade é fundamental para o sucesso de qualquer empresa (Caro e Garcia, 2007) que deseja fidelizar e adquirir clientes. Portanto, compreender a satisfação do cliente é essencial para as empresas que desejam prosperar no mercado online.

As informações sobre a satisfação do cliente em plataformas eletrônicas de comércio são obtidas de diversas maneiras, incluindo a coleta de feedback por meio de formulários de pesquisa, pesquisas por e-mail após uma compra, mensagens de SMS, comentários em redes sociais, avaliações em plataformas de avaliação de produtos e serviços e outras formas (HubSpot, 2022) como até o monitoramento de redes sociais para que as empresas possam avaliar a recepção de seus produtos, a qualidade de seus serviços e a percepção de sua imagem para o público.

Em uma parcela desses questionários de pesquisa de satisfação, é utilizada a escala de Likert, um modelo de verificação para “*mensurar atitudes no contexto das ciências comportamentais*”, que é adequada para instrumentos de pesquisa longos com dados e conclusões confiáveis, medindo o nível de concordância através de uma escala de números (Silva Júnior & Costa, 2014) em perguntas como “O quão satisfeito você ficou com o nosso atendimento?”. Além disso, ocasionalmente é pedido que o cliente elabore respostas mais livres para perguntas como “O que poderíamos ter feito melhor?” para uma percepção mais clara do problema. Já em plataformas de avaliação de produtos e serviços, como o Reclame Aqui¹, o cliente pode realizar reclamações referentes aos serviços e produtos de diversas empresas, e as mesmas ainda podem respondê-las.

Tais informações ricas e variadas são essenciais para compreender como os clientes percebem suas interações com as empresas de e-commerce. Segundo Turchi (2012), “*É fundamental conhecer muito bem o seu público e seus interesses, pois é a única maneira de oferecer conteúdo relevante e de qualidade a ele*”.

As empresas implementam uma variedade de abordagens para medir a satisfação do cliente. Isso pode envolver o uso de métricas tradicionais recorrentemente usadas, como o Net Promoter Score (NPS), o Customer Satisfaction Score (CSAT) e o Customer Effort Score (CES) onde, respectivamente, busca compreender a fidelidade do consumidor, satisfação dos usuários e experiência com o serviço (Silva, 2022). Além disso, para que seja possível melhorar o relacionamento com seus clientes através da coleta, armazenamento e análise de informações sobre clientes, algumas empresas fazem uso de sistemas de Customer Relationship Management (CRM) para estimular a fidelização dos mesmos (BigDataCorp, 2023).

¹ Reclame Aqui. Disponível em: <https://www.reclameaqui.com.br/>

No entanto, a análise de dados por mineração de texto emerge como uma ferramenta que gera resultados inestimáveis neste contexto. Através da mineração de texto as empresas podem extrair insights significativos a partir de dados não estruturados, como os comentários dos clientes em redes sociais e *reviews* de produtos ou serviços, trazendo diversos benefícios importantes às empresas (Support Logic, 2022) que desejam impulsionar seus negócios.

A mineração de texto é definida por técnicas de extração de padrões ou tendências de grandes volumes textuais que resultam na descoberta de anomalias e padrões valiosos (Barion & Lago, 2008). No contexto da satisfação do cliente, a análise de sentimentos é uma técnica central, aplicada para determinar os principais destaques relacionados à satisfação do cliente em relação aos seus aspectos (Hong et. al., 2019). Tal análise utiliza técnicas de processamento de linguagem natural e computação para identificar automaticamente e classificar o sentimento de textos (Hussein, 2018). Por exemplo, em uma análise de sentimentos feita com consumidores de produtos orgânicos foram descobertos diversos fatores que incentivam os clientes a realizarem suas compras, assim como os descontos dos preços dos produtos vendidos online, a qualidade da comida, as informações nutritivas e entre outros, assim como foram descobertos fatores que desestimulam tais compras (Lyu & Choi, 2020).

Outra técnica utilizada é a modelagem de tópicos, técnica estatística composta por algoritmos para revelar e descrever temas em uma coleção de documentos (Kherwa & Bansal, 2019), onde método mais usado é o Latent Dirichlet Allocation (LDA) que, através do mesmo, é possível descobrir palavras representativas de um ou mais tópicos em documentos textuais (Blei et. al., 2003). Para que fosse possível expandir o entendimento do problema, na mesma pesquisa sobre consumidores de produtos orgânicos no e-commerce também foi utilizada tal técnica para relacionar os tópicos obtidos com os sentimentos analisados (Lyu & Choi, 2020). Dessa forma, foi estudada a relação entre ambos para um melhor entendimento sobre a satisfação dos clientes.

3. OBJETIVO E PROPOSTA DO TRABALHO

Este projeto visa aprimorar significativamente a identificação da satisfação do consumidor no cenário do comércio eletrônico brasileiro. Para alcançar esse objetivo,

aplicaremos técnicas avançadas de mineração de texto em comentários e avaliações pós compra de serviços e produtos em plataformas de e-commerce. Os resultados esperados incluem a capacidade de discernir não apenas se os clientes estão satisfeitos ou insatisfeitos, mas também compreender as relações entre os sentimentos e as principais nuances de suas opiniões. Isso permitirá uma análise mais abrangente e precisa, resultando em insights valiosos para aprimorar a qualidade dos produtos, serviços e experiências de compra.

4. PLANO DE AÇÃO

Este projeto divide-se nas seguintes etapas, conforme demonstrado na Figura 1:

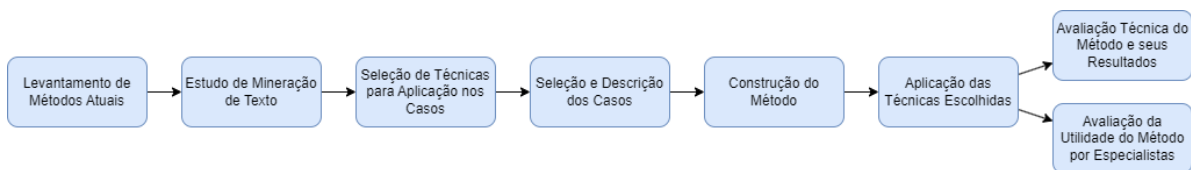


Figura 1 - Etapas do plano de ação. Fonte: A autora.

- 1) Levantamento de Métodos Atuais: Pesquisa dos métodos existentes para avaliação de satisfação do cliente.
- 2) Estudo de Mineração de Texto: Investigação sobre técnicas de mineração de texto com propósito de adquirir um conhecimento sólido de métodos e ferramentas.
- 3) Seleção de Técnicas para Aplicação nos Casos: Escolha das técnicas mais apropriadas para aplicação em um caso de análise de satisfação do cliente.
- 4) Seleção e Descrição dos Casos: Escolha de casos representativos e pertinentes para a aplicação das técnicas selecionadas.
- 5) Construção do Método: Estruturamento e consolidação do método, contemplando a aplicação das técnicas selecionadas.
- 6) Aplicação das Técnicas Escolhidas: Implementação das técnicas previamente selecionadas aos dados coletados.
- 7) Avaliação Técnica do Método e seus Resultados: Inspeção técnica dos resultados obtidos de forma a garantir a qualidade da análise.
- 8) Avaliação da Utilidade do Método por Especialistas: Inspeção da utilidade prática dos resultados por meio de entrevistas com especialistas em análise de dados.

Cronograma previsto

Etapas	Out	Nov	Dez	Jan	Fev	Mar	Abr	Mai	Jun
Etapa 1									
Etapa 2									
Etapa 3									
Etapa 4									
Etapa 5									
Etapa 6									
Etapa 7									
Etapa 8									

Cronograma realizado

Etapas	Out	Nov	Dez	Jan	Fev	Mar	Abr	Mai	Jun
Etapa 1									
Etapa 2									
Etapa 3									
Etapa 4									
Etapa 5									
Etapa 6									
Etapa 7									
Etapa 8									

5. LEVANTAMENTO DOS MÉTODOS ATUAIS

A coleta de feedback dos clientes desempenha um papel fundamental na compreensão das opiniões e experiências dos consumidores, proporcionando às empresas informações valiosas para aprimorar produtos, serviços e a experiência global do cliente. Tal coleta pode ser conduzida de duas maneiras distintas, dependendo de quem assume o papel ativo: o cliente ou a própria empresa. Com isso

em mente, foram definidas duas abordagens predominantes para a coleta de feedback dos clientes: as abordagens espontânea e por demanda. Cada uma delas possui características próprias, conforme ilustrado na Figura 2, e métodos específicos de implementação e foram pensadas diretamente pelo autor com base em artigos.

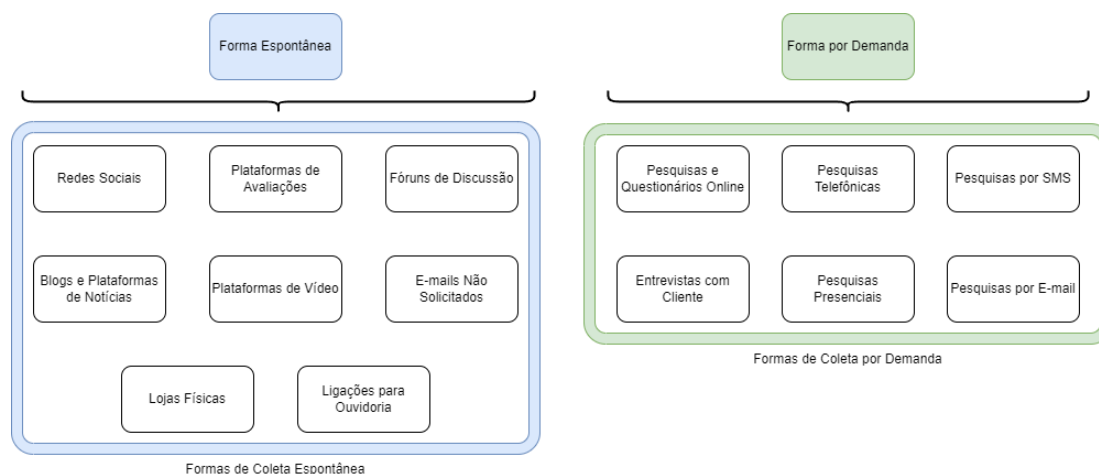


Figura 2 - Diferenças entre a forma espontânea e por demanda de coleta de feedback.
Fonte: A autora.

A coleta espontânea de feedback do cliente é caracterizada pela iniciativa do próprio cliente em compartilhar suas opiniões e experiências. Essa abordagem não envolve uma solicitação direta por parte da empresa e as informações são obtidas de forma voluntária, conforme o seguinte:

- Comentários em redes sociais: Os clientes compartilham suas opiniões nas páginas e postagens das empresas nas redes sociais, sem que haja solicitação ativa da empresa.
- Plataformas de avaliação: Sites como Reclame Aqui e Google Avaliações² permitem que os clientes expressem suas experiências e avaliações sem que a empresa os convide a fazê-lo.
- Comentários em blogs e fóruns: Os clientes podem deixar comentários espontâneos em blogs, fóruns de discussão e plataformas de notícias relacionados à empresa ou aos produtos.
- Feedback em vídeos: Comentários em vídeos publicados por empresas ou

² O Google Avaliações é um recurso integrado ao Google Maps que permite aos usuários compartilharem suas opiniões sobre locais comerciais. Consulte <https://support.google.com/maps/answer/6230175> para mais informações.

youtubers sobre produtos ou serviços.

- E-mails não solicitados: Alguns clientes enviam e-mails com feedback não solicitado, compartilhando suas experiências e opiniões diretamente com a empresa.
- Avaliações em lojas físicas: Em lojas físicas, os clientes podem compartilhar feedback por meio de comentários e sugestões em formulários ou diretamente com a equipe de atendimento.
- Ligações para ouvidoria: As ligações diretas para a ouvidoria da empresa onde os clientes expressam suas opiniões e preocupações.

A coleta por demanda de feedback do cliente é caracterizada pela iniciativa da empresa em buscar a opinião dos consumidores de forma proativa e estruturada. Nesse caso, a empresa solicita informações específicas da clientela para avaliar a satisfação e experiência, conforme o seguinte:

- Pesquisas por SMS: As empresas enviam pesquisas de satisfação por meio de mensagens de texto solicitando que os clientes classifiquem suas experiências.
- Pesquisas e questionários online: Pesquisas online por e-mail ou formulários em sites são comuns para avaliar a satisfação do cliente.
- Pesquisas telefônicas: Além das pesquisas online, algumas empresas realizam pesquisas por telefone, coletando feedback diretamente do consumidor.
- Entrevistas com o cliente: As entrevistas individuais com os clientes permitem uma compreensão mais aprofundada das experiências e opiniões.
- Pesquisas presenciais: Em configurações físicas, como lojas ou estabelecimentos de atendimento ao cliente, as empresas realizam pesquisas presenciais.
- Pesquisas por e-mail: As pesquisas por e-mail são frequentemente usadas para avaliar a satisfação do cliente após interações específicas.

Além disso, na coleta por demanda, as empresas frequentemente empregam métricas de avaliação para quantificar a satisfação e o esforço do cliente quando são realizadas pesquisas através de formulários estruturados. Três métricas comuns incluem:

- Net Promoter Score (NPS): O NPS avalia a probabilidade de um cliente

recomendar a empresa a outras pessoas, fornecendo uma métrica padronizada para medir a lealdade do cliente (Weber, 2021).

- Customer Satisfaction Score (CSAT): O CSAT mede a satisfação do cliente com uma interação específica, produto ou serviço. Os clientes geralmente são solicitados a classificar sua satisfação em uma escala de 1 a 5 ou de 1 a 7, com 1 sendo “muito insatisfeito” e 7 sendo “muito satisfeito” (Weber, 2021).
- Customer Effort Score (CES): O CES mede o esforço que os clientes precisam fazer para resolver um problema ou alcançar um objetivo com a empresa. Os clientes são solicitados a classificar o esforço em uma escala de fácil a difícil (Silva, 2022).

6. ESTUDO DE MINERAÇÃO DE TEXTO

A mineração de texto desempenha um papel crucial na ciência de dados, concentrando-se na identificação de padrões e relações ocultas presentes em dados não estruturados, como documentos de texto. Hotho, Nürnberger e Paaß (2005) destacam a integração de técnicas de recuperação de informação, processamento de linguagem natural (PNL), além de métodos de *Knowledge Discovery in Databases* (KDD), mineração de dados, aprendizado de máquina e estatísticas nesse domínio.

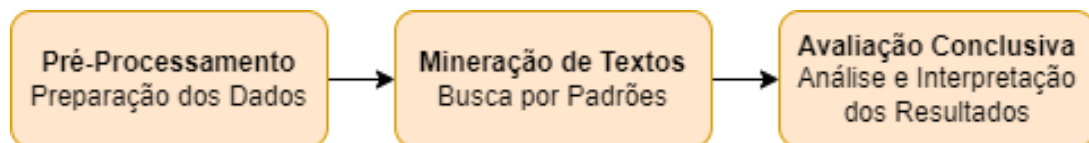


Figura 3 - Processo de mineração de texto, baseado em Maciel et. al. (2015). Fonte: A autora.

O processo de mineração de texto tem seu ponto de partida na aquisição de dados brutos que podem ser provenientes de diversas fontes como redes sociais, documentos, artigos de notícias, e-mails ou qualquer outra fonte de texto não estruturado. Após a coleta, entra em cena o pré-processamento, um estágio crucial que adapta os dados para uma análise eficaz. Tal etapa envolve a limpeza do conteúdo textual onde é feita a normalização, que visa empregar uma padronização do conteúdo textual assim como converter todas as palavras em letras minúsculas, e remoção de pontuações e palavras sem importância, conforme abordado por Kalra e Aggarwal (2017). Com essas etapas concluídas, o conteúdo textual está pronto para

prosseguir para análises mais aprofundadas utilizando práticas de classificação e extração de texto, ou seja, a própria mineração de dados, até ser possível realizar a interpretação dos dados.

6.1. PRÉ-PROCESSAMENTO

O pré-processamento textual, em termos simples, significa organizar os dados em um formato previsível e analisável. Tal passo é essencial para preparar o “terreno” antes de aplicar as técnicas de mineração de texto já que tal prática não apenas assegura uma limpeza adequada dos dados, mas também se concentra na normalização dos dados, como foi dito anteriormente. A garantia de uma base de dados sólida nesta fase inicial contribui significativamente para resultados mais precisos nas etapas subsequentes da análise de texto (Yang, 2008).

Após a realização da limpeza do texto, ou seja, remoção de pontuações, caracteres e palavras sem importância, há diversas técnicas que são comumente implementadas no pré-processamento dos dados como, por exemplo, a tokenização, a remoção de *stopwords* e o *stemming* ou lematização, dependendo do contexto da aplicação. A tokenização divide o texto em unidades menores (tokens) enquanto o *stemming* e a lematização reduzem as palavras à sua forma base (Patel, Passi, 2020).

Com base nas informações e explicações fornecidas por Silva (2023) o pré-processamento de texto deve passar por alguns processos fundamentais. No estágio inicial, a tokenização desempenha um papel crucial, fragmentando um bloco de texto como, por exemplo, “Eu vejo um copo de café” em unidades menores, resultando no conjunto [“Eu”, “vejo”, “um”, “copo”, “de”, “café”].

A subsequente remoção de *stopwords* tem como objetivo eliminar palavras comuns tais como artigos e preposições que geralmente não contribuem significativamente para a análise de texto. Por exemplo, ao retirar as *stopwords* do texto [“Esse”, “é”, “um”, “exemplo”, “de”, “remoção”, “de”, “stopwords”], obtemos uma estrutura mais focada e relevante para análises posteriores: [“Esse”, “exemplo”, “remoção”, “stopwords”].

No contexto do *stemming*, ocorre a redução de palavras à sua forma base, ou seja, ao seu radical. Por exemplo, a aplicação do *stemming* transformaria as palavras “corrida” e “correndo” para a base “corr”, simplificando a análise ao agrupar variações

de uma palavra e tornando a compreensão do texto mais eficiente.

A lematização, por sua vez, compartilha semelhanças com o stemming, mas adota uma abordagem mais refinada reduzindo as palavras à sua forma no dicionário, mantendo o significado original. Nesse caso, palavras como “chuvisqueiro” e “chuvisco” seriam transformadas para forma base “chuva”. Tal técnica é útil quando é crucial manter a integridade semântica do texto.

Com essas etapas concluídas, o conteúdo textual está pronto para prosseguir para análises mais aprofundadas utilizando práticas de classificação e extração de texto.

6.2. MINERAÇÃO DE TEXTO

A classificação de texto, conforme MonkeyLearn (2023), é uma prática essencial na mineração de texto que emprega algoritmos de aprendizado de máquina para atribuir categorias a dados não estruturados. Treinados com exemplos categorizados, tais algoritmos demonstram a capacidade de classificar novos textos com precisão após o treinamento. Como já foi explicado anteriormente, essa abordagem abrange diversas aplicações, tais como a análise de tópicos para identificação dos temas principais e a análise de sentimentos para a compreensão das emoções subjacentes em um texto.

A detecção do idioma é outra aplicação crucial da classificação de texto. Essa técnica permite a classificação automática de textos com base em sua língua (Rajanak, Patil, Singh, 2023), facilitando a organização e a interpretação de grandes volumes de dados multilíngues. Tal capacidade é comumente usada para a implementação de tradução automática, que busca traduzir o conteúdo automaticamente de uma língua para outra, mantendo o significado original. O Google Translate³ é um dos exemplos de ferramentas hodiernas que utilizam tal funcionalidade, como destacado por Qizi Samiyeva (2023). Além disso, no cenário das avaliações de clientes sobre serviços ou produtos, tal detecção pode também ser útil para filtrar apenas os textos escritos em língua portuguesa.

A detecção de intenção emerge como uma outra técnica na classificação de texto. Tal abordagem tem como objetivo automatizar o reconhecimento dos propósitos

³ Google Translate. Disponível em: <https://translate.google.com.br/>.

subjacentes em um texto, desvendando a intenção por trás das palavras utilizadas. A aplicação dessa técnica foi examinada em um estudo conduzido por Akulick (2017) concentrando-se exclusivamente na identificação das intenções dos clientes em relação a avaliações sobre produtos e serviços. O estudo abordou categorias específicas tais como intenções de compra ou desistência, recomendação ou aviso, e elogio ou reclamação.

Paralelamente, a extração de texto emerge como outra prática crucial na obtenção de informações específicas, conforme destacado pela IBM (2023). Tal método abrange diversas aplicações como a seleção de recursos, a extração de características e o reconhecimento de entidades nomeadas. A seleção de recursos, ou seleção de atributos, é o processo de escolha dos recursos (dimensões) mais importantes que contribuem ao máximo para a saída de um modelo de análise preditiva. Já a extração de recursos envolve a escolha de um subconjunto de recursos que melhora a precisão de uma tarefa de classificação, sendo especialmente relevante para a redução da dimensionalidade. Essas técnicas, fundamentais no processamento de linguagem natural, convergem para aprimorar a mineração de texto, permitindo a extração de dados importantes.

A seleção de recursos é uma técnica que se destaca pela redução das entradas para o processamento e análise, ou de encontrar as entradas mais significativas, escolhendo um pequeno subconjunto de recursos relevantes dos originais, removendo ruídos e irrelevâncias (Miao, Niu, 2016). Ou seja, o objetivo é maximizar a relevância e minimizar a redundância ao encontrar um subconjunto que consiste apenas em recursos relevantes (Jović, Brkić, Bogunović, 2015). Na perspectiva da aplicação, o método é utilizado em alguns casos de categorização de texto, de recuperação de informações musicais, de recuperação de imagens e até de análise genômica (Kumar, Minz, 2014).

A extração de características representa uma das técnicas importantes de redução de dimensionalidade usadas no aprendizado de máquina para mapeamento de dados de grandes dimensões, aumentando o desempenho dos modelos e reduzindo a complexidade computacional (Avinash, Sivasankar, 2019). Tal método é muito utilizado em processos de visão computacional, recuperação de imagens, reconhecimento de fala, bioinformática e entre outros, sendo muito útil na mineração de texto para recuperação de informações (Salau, Jain, 2019). Adicionalmente,

destaca-se a extração de características de produtos que, em uma pesquisa conduzida por Wei (2010) foi possível compreender as nuances das opiniões dos compradores online.

Por fim, o reconhecimento de entidades nomeadas é uma técnica que desempenha um papel significativo na localização e categorização de substantivos importantes e nomes próprios em um texto. Tomemos como exemplo a frase: “Antes de ingressar na UCB, Lisa North trabalhou para a Pegasus Books em North Berkeley”. Nesse contexto, seriam identificadas a UCB e Pegasus Books como organizações, Lisa North como uma pessoa e North Berkeley” como um local (Mohit, 2014). Em determinadas circunstâncias, como ilustrado na pesquisa de Zhao e Liu (2008), essa técnica pode ser aplicada à extração de entidades nomeadas relacionadas a produtos, facilitando a identificação de seus respectivos nomes e marcas.

6.3. AVALIAÇÃO CONCLUSIVA

Finalmente, após a aplicação das diversas técnicas de mineração de texto, a etapa crucial de pós-processamento emerge como o último elo na cadeia da análise. Nessa fase, nos concentramos na interpretação dos resultados obtidos. Os insights extraídos, categorizações realizadas e padrões identificados são examinados para derivação de conclusões. Esse estágio não apenas valida a eficácia das técnicas empregadas mas também oferece uma visão abrangente do contexto, possibilitando a compreensão do impacto prático dos resultados.

7. SELEÇÃO DE TÉCNICAS PARA APLICAÇÃO NOS CASOS

As técnicas selecionadas a serem aplicadas nos casos de uso foram, em ordem cronológica de execução, a modelagem de tópicos, a análise de sentimentos por tópico e a extração de entidades nomeadas por tópicos a fim de sintetizar os resultados de acordo com os temas contidos nos comentários das avaliações. A modelagem de tópicos permite identificar tais assuntos principais discutidos pelos consumidores, enquanto a análise de sentimentos por tópicos oferece uma visão detalhada das emoções e opiniões associadas a cada tema. A extração de entidades nomeadas por tópico, por fim, destacam as entidades específicas mencionadas

igualmente de acordo com a temática.

A escolha de utilizar análise de tópicos, análise de sentimentos e reconhecimento de entidades nomeadas para avaliar a satisfação dos clientes em relação a serviços e produtos no contexto do e-commerce através de suas avaliações é fundamentada na necessidade de extração de insights significativos de volumes de dados textuais em língua portuguesa. Essas técnicas representam pilares importantes na mineração de texto, proporcionando uma compreensão aprofundada das nuances presentes nos comentários dos consumidores. Dessa forma, a aplicação dessa abordagem visa contribuir de maneira significativa para o aprimoramento contínuo de produtos, serviços e atendimentos pela percepção as principais problemáticas.

A análise de tópicos desempenha um papel crucial ao identificar e categorizar os principais temas abordados nas avaliações, proporcionando uma visão sistêmica dos assuntos mais discutidos pelos usuários. Isso possibilita uma compreensão ampla dos interesses e preocupações recorrentes. Por exemplo, ao analisar as avaliações de um serviço ou produto é possível discernir quais são as reclamações e elogios mais recorrentes. Além disso, é viável destacar informações de tópicos específicos como tempo de entrega, qualidade de produto, facilidade de uso, entre outros.

Adicionalmente, a análise de sentimentos é fundamental para capturar a carga emocional por trás das avaliações, permitindo avaliar a satisfação geral dos clientes e identificar áreas específicas que podem exigir melhorias. Por exemplo, ao examinar as opiniões de clientes sobre um serviço de entrega poderíamos identificar se os comentários estão predominantemente satisfeitos com a rapidez e eficiência do serviço ou se há insatisfações relacionadas a atrasos ou problemas de entrega.

O reconhecimento de entidade nomeada complementa esse conjunto de técnicas ao permitir a identificação de organizações, pessoas e locais mencionados nas avaliações. Esse processo é particularmente valioso para entender quem ou o que está associado a determinados sentimentos ou tópicos. Por exemplo, ao examinar as avaliações de um restaurante, seria possível identificar se as críticas negativas estão concentradas em um prato específico ou em algum indivíduo, como o chefe de cozinha ou garçom.

Em conjunto, essas técnicas oferecem uma abordagem holística para a análise de avaliações de produtos e serviços online, proporcionando não apenas uma compreensão profunda das opiniões dos consumidores, mas também a extração de

informações valiosas que podem orientar estratégias de negócios, melhorias de produtos e aprimoramentos nos serviços prestados. Essa escolha visa maximizar a utilidade e a eficácia da análise de dados textuais no contexto específico de avaliações online.

A decisão de não focar, neste momento, na extração de características específicas de produtos foi tomada de forma deliberada, já que neste trabalho o foco é na realização de uma análise mais abrangente e generalizada das avaliações de produtos e serviços. Além disso, atualmente, os métodos disponíveis para essa técnica apresentam diversas formas de aplicação onde, a maioria delas, implementa o uso de técnicas de análise de tópicos. Em um dos casos, foi utilizado o algoritmo Latent Dirichlet Allocation (LDA) para modelagem de tópicos com o reconhecimento de sinônimos para identificar similaridades entre produtos (Feng et al., 2019). Como nosso objetivo não é analisar individualmente os produtos e serviços, focaremos na abordagem da extração de tópicos.

8. SELEÇÃO E DESCRIÇÃO DOS CASOS

Conforme mencionado anteriormente, o propósito deste trabalho é o desenvolvimento de uma metodologia para avaliação da satisfação dos clientes em plataformas de comércio eletrônico brasileiros, utilizando os comentários dos consumidores sobre produtos e serviços como o principal foco do projeto. Neste contexto, a obtenção de um conjunto de dados adequado tornou-se crucial para a viabilização da modelagem de tópicos, análise de sentimentos e extração de entidades nomeadas.

Para assegurar a profundidade e abrangência da pesquisa, a busca por um conjunto de dados com informações verídicas de tais comentários em língua portuguesa tornou-se uma prioridade. A coleta de satisfações de forma ativa requer o acesso à informações ocasionalmente privadas de negócios e, portanto, foi preferida a procura por opiniões obtidas de forma espontânea. Diante disso, foram cogitadas duas opções para a obtenção de tais dados, onde a primeira se basearia na realização de um processo de web scraping, que se fundamenta na extração de informações diretamente de sites, e na procura por bases de dados públicas.

O objetivo era obter um conjunto de dados que não fosse pequeno, ou seja, com

pelo menos mais de 1.000 avaliações e, por conta disso, a opção da realização de um web scraping de uma plataforma de comércio eletrônico foi descartada, considerando o tempo significativo que essa abordagem demandaria. Vale ressaltar que tal alternativa foi implementada e testada baseando-se na utilização do XPath no HTML da página de um único restaurante no TripAdvisor⁴ para coletar exclusivamente os comentários e a quantidade de estrelas das avaliações do mesmo. No entanto, essa estratégia revelou-se demorada, consumindo cerca de 5 horas em meu computador pessoal para a coleta de cerca de 2.000 comentários. Logo, a escolha foi direcionada para a busca de bases de dados públicas disponíveis na internet.

Inicialmente, a busca por conjuntos de dados públicos em português levou-me ao Kaggle⁵, uma plataforma reconhecida por suas competições de ciência de dados e recursos para entusiastas da área. No âmbito dessa busca, encontrei um conjunto fornecido pela Olist⁶, uma plataforma de vendas digital presente desde 2015 nos principais marketplaces do Brasil. Tal conjunto abrange informações detalhadas sobre compras de produtos, incluindo perfis de consumidores, dados de pagamentos, informações sobre os produtos e outros aspectos relacionados às compras. Com cerca de suas 100.000 linhas de dados datados de 2016 a 2018, juntamente à confiabilidade da fonte, ele tornou-se minha primeira escolha de fonte para o projeto.

Na continuidade da busca por um conjunto de dados, constatei que a maioria deles carecia de informações em língua portuguesa ou não se alinhavam ao contexto do comércio eletrônico. Diante desse cenário, ampliei a pesquisa utilizando uma ferramenta dedicada à busca de conjuntos de dados públicos: o Dataset Search⁷ da Google. Nessa abordagem, deparei-me com um conjunto contendo avaliações de consumidores da B2W Digital⁸, uma plataforma de comércio eletrônico criado pela união da Americanas⁹ com a Submarino¹⁰, contendo cerca de 130.000 linhas de dados datados de Janeiro a Maio de 2018, disponibilizado pelo repositório da Americanas no GitHub. Tal conjunto engloba informações detalhadas sobre os consumidores, incluindo sua geolocalização, bem como os produtos que foram

⁴ TripAdvisor. Disponível em: <https://www.tripadvisor.com.br/>

⁵ Kaggle. Disponível em: <https://www.kaggle.com/>

⁶ Olist. Disponível em: <https://olist.com/>

⁷ Dataset Search. Disponível em: <https://datasetsearch.research.google.com/>

⁸ B2W Digital. Disponível em: <https://en.wikipedia.org/wiki/B2W>

⁹ Americanas. Disponível em: <https://www.americanas.com.br/>

¹⁰ Submarino. Disponível em: <https://www.submarino.com.br/>

avaliados. Assim, surgiu mais uma fonte de dados relevante para o trabalho.

Contudo, à medida que a procura prosseguia por mais conjuntos para avaliar se não seriam encontradas outras opções interessantes, deparei-me novamente com a mesma situação do Kaggle, onde a busca por outros conjuntos em língua portuguesa que atendessem às minhas necessidades não obteve sucesso. Diante dessa situação, o foco foi direcionado para o Google Acadêmico¹¹ em busca de trabalhos que tivessem utilizado conjuntos de dados com avaliações de produtos ou serviços. Após certo tempo, curiosamente, foi encontrado um estudo brasileiro de análise de sentimentos que não apenas utilizou os conjuntos de dados da Olist e da B2W Digital, mas também outros dois conjuntos: da Buscapé¹², uma plataforma de comércio eletrônico brasileiro, com cerca de 80.000 linhas de dados datados de 2013 e UTLC, com cerca de 2.000.000 de linhas de dados datados de 2019, sendo esse último dividido em comentários de filmes retirados de um aplicativo chamado Filmow¹³ e, o outro, de comentários sobre aplicativos da Play Store¹⁴. Apesar da referência original a esses dois últimos conjuntos estar indisponível, uma investigação mais aprofundada revelou que um usuário no Kaggle havia compartilhado os quatro conjuntos, contendo dados provenientes de comentários e a quantidade de estrelas, indo de 1 a 5, das avaliações.

Vale ressaltar que os conjuntos de dados da Buscapé e da UTLC foram coletados por meio do projeto Opinando, uma iniciativa da Universidade de São Paulo (USP). O propósito desse projeto foi conduzir uma mineração de opiniões para realização de análises e sínteses do conteúdo textual presente nas avaliações.

Durante o procedimento de aplicação do método nos conjuntos de dados, foi notada uma necessidade, que será explicada posteriormente, que resultou na descoberta, através do Dataset Search, de um conjunto composto por avaliações de hotéis na plataforma do TripAdvisor que foram coletadas de Fevereiro até Março de 2018 de todos os estados brasileiros.

Por fim, pelo objetivo do trabalho ser analisar avaliações de usuários do comércio eletrônico brasileiro, foi optada a utilização de quatro conjuntos: da Olist, da Buscapé,

¹¹ Google Acadêmico. Disponível em: <https://scholar.google.com.br/>

¹² Buscapé. Disponível em: <https://www.buscape.com.br/>

¹³ Filmow. Disponível em: <https://filmow.com/>

¹⁴ Play Store. Disponível em: <https://play.google.com/store/>

da B2W Digital e de hotéis brasileiros. A escolha desses quatro conjuntos de dados específicos se justifica pela ampliação da cobertura do estudo ao combinar os resultados dessas avaliações de diferentes fontes, capturando uma visão mais abrangente e particular de cada plataforma. Além disso, considerando o tempo e recursos, a escolha de quatro conjuntos oferece um equilíbrio entre a abrangência e a viabilidade prática para execução dos estudos.

O foco central estará voltado para o conteúdo textual dos comentários dos consumidores e nas avaliações em estrelas, sendo esse último utilizado para a execução do treinamento de um modelo de linguagem que foi utilizado para realização da análise de sentimentos, visando proporcionar aos gestores de negócios uma análise abrangente que inclui tópicos, sentimentos e identificação de entidades nomeadas. Um modelo de linguagem é um sistema de inteligência artificial que usa aprendizado de máquina para prever palavras com base no contexto, sendo útil para tradução automática e geração de texto (Wikipédia). O intuito é fornecer uma simples compreensão dos aspectos positivos e negativos de seus comércios eletrônicos.

A seguir, são apresentadas as fichas com informações gerais de cada conjunto.

	GUACAMOLE
Origem	Restaurante Guacamole no TripAdvisor
Tamanho	2.872 avaliações utilizadas
Período	Avaliações até Dezembro de 2023
Formato	Arquivo CSV
Colunas principais	Data da avaliação, comentário da avaliação
URL da fonte	https://www.tripadvisor.com.br/Restaurant_Review-g303506-d3399400-Reviews-Guacamole_Cocina_Mexicana_Barra_da_Tijuca-Rio_de_Janeiro_State_of_Rio_de_Janeiro.html

	OLIST
Origem	Plataforma de comércio eletrônico Olist

Tamanho	37.564 avaliações utilizadas
Período	Avaliações de 2016 até 2018
Formato	Arquivo CSV
Colunas principais	Data da avaliação, comentário da avaliação, estrelas da avaliação, valor do pagamento do pedido, cidade do cliente, entre outras
URL da fonte	https://www.kaggle.com/datasets/olistbr/brazilian-ecommerce

	BUSCAPÉ
Origem	Plataforma de comércio eletrônico Buscapé
Tamanho	79.025 avaliações utilizadas
Período	Avaliações de 2013
Formato	Arquivo CSV
Colunas principais	Comentário da avaliação, estrelas da avaliação, recomendação a um amigo
URL da fonte	https://www.kaggle.com/datasets/fredericods/ptbr-sentiment-analysis-datasets

	B2W DIGITAL
Origem	Plataforma de comércio eletrônico B2W Digital
Tamanho	129.863 avaliações utilizadas
Período	Avaliações de Janeiro a Maio de 2018
Formato	Arquivo CSV
Colunas principais	Comentário da avaliação, estrelas da avaliação, título da avaliação, nome do produto, entre outras
URL da fonte	https://github.com/americanas-tech/b2w-reviews01

	HOTÉIS BRASILEIROS
Origem	Hotéis brasileiros no TripAdvisor
Tamanho	686.741 avaliações utilizadas
Período	Avaliações de Fevereiro a Março de 2018
Formato	Arquivo CSV
Colunas principais	Comentário da avaliação, estrelas da avaliação, título da avaliação e data da avaliação
URL da fonte	https://data.mendeley.com/datasets/2w3kvrg97m/1

	HOTEL GRAND HYATT
Origem	Hotel Grand Hyatt de São Paulo no TripAdvisor
Tamanho	2.623 avaliações utilizadas
Período	Avaliações de Fevereiro a Março de 2018
Formato	Arquivo CSV
Colunas principais	Comentário da avaliação, estrelas da avaliação, título da avaliação e data da avaliação
URL da fonte	https://data.mendeley.com/datasets/2w3kvrg97m/1

	SMARTPHONE MOTOROLA
Origem	Smartphone Motorola da plataforma de comércio eletrônico B2W
Tamanho	792 avaliações utilizadas
Período	Avaliações de Janeiro a Maio de 2018
Formato	Arquivo CSV
Colunas principais	Comentário da avaliação, estrelas da avaliação, título da avaliação, nome do produto, entre outras

URL da fonte	https://github.com/americanas-tech/b2w-reviews01
--------------	---

9. CONSTRUÇÃO DO MÉTODO

Para a construção do método, as técnicas de mineração de texto foram inicialmente aplicadas no conjunto de dados do restaurante Guacamole, obtidos por meio da plataforma do TripAdvisor em Dezembro de 2023, utilizando a técnica de web scraping. Composto por cerca de 2.800 avaliações, tal conjunto foi selecionado para um primeiro estudo das técnicas por conta de sua extensão, o que conseguiu agilizar a análise e simplificar a compreensão da implementação e dos resultados nas aplicações.

Posteriormente, a aplicação do método expandiu-se para os quatro conjuntos de dados escolhidos anteriormente: nas avaliações da Olist, com cerca de 100.000 comentários, da Buscapé, com cerca de 80.000 comentários, da B2W Digital, com cerca de 130.000 comentários, e de hotéis brasileiros, com cerca de 730.000 comentários. O uso deste último conjunto, em específico, será apresentado e esclarecido mais à frente.

Por conta do seu uso em diversas pesquisas em português, foi escolhida a utilização de técnicas e modelos baseados no modelo BERT de inteligência artificial da Google, como o BERTopic¹⁵ para a modelagem de tópicos, que utiliza a técnica TF-IDF (Term Frequency-Inverse Document Frequency) para extração dos termos mais relevantes do conteúdo textual, e o BERTimbau¹⁶ para a análise de sentimentos. Além da robustez e eficácia comprovadas do BERT em tarefas de processamento de linguagem natural, essas técnicas e modelos específicos oferecem vantagens adicionais, como a capacidade de lidar com textos em português de forma eficiente e a integração de técnicas avançadas de modelagem de tópicos e análise de sentimentos.

O pré-processamento dos comentários das avaliações para a modelagem de tópicos foi feito da mesma forma para todos os conjuntos de dados, começando com a etapa de limpeza textual de tais comentários. Nessa fase, foram removidos dígitos, pontuações, quebras de linha e palavras curtas com até 2 letras consideradas pouco

¹⁵ BERTopic. Disponível em: <https://github.com/MaartenGr/BERTopic>

¹⁶ BERTimbau. Disponível em: <https://neuralmind.ai/bert/>

relevantes para a análise, sendo tudo realizado por meio do uso de expressões regulares. Em seguida, foi aplicada a técnica de tokenização para segmentar o texto em palavras individuais, o que possibilitou o processo de remoção das stopwords em língua portuguesa.

Após a eliminação das stopwords, iniciou-se a lematização. O objetivo dessa etapa foi reduzir a dimensionalidade do vocabulário e melhorar a consistência dos dados, mantendo apenas os substantivos para uma compreensão mais clara dos principais tópicos abordados nas avaliações. Por fim, foi realizada a remoção de acentuação das palavras, buscando padronizar o texto e facilitar a comparação entre palavras acentuadas e não acentuadas já que a forma em que as palavras são escritas ocasionalmente são diferentes nos comentários.

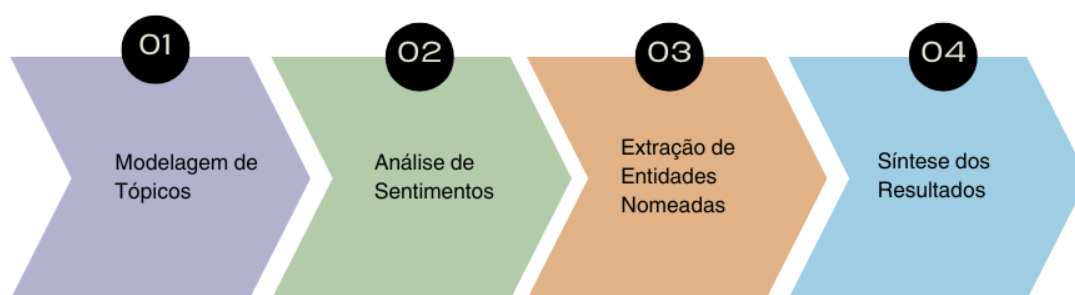


Figura 4 - Processos do método consolidado. Fonte: A autora.

Com o pré-processamento concluído, é possível avançar para a etapa da aplicação da mineração de texto. Portanto, foi escolhido iniciar o processo com a modelagem de tópicos para prosseguir posteriormente com a análise de sentimentos e a extração de entidades nomeadas. O motivo dessa ordem de preferência de processamento se deve ao estabelecimento de uma organização progressiva na análise dos comentários já que, ao identificar, ajustar e compreender os tópicos principais abordados nas avaliações, é possível direcionar de forma mais eficaz as etapas subsequentes da análise.

A implementação da modelagem de tópicos utilizando o BERTopic foi conduzida de forma padronizada para o conjunto de dados do restaurante Guacamole, sem modificar os parâmetros das cinco etapas independentes: embedding, redução de dimensionalidade, clusterização, tokenização e esquema de ponderação. No entanto, ao lidar com outros conjuntos de dados, os parâmetros dessas etapas foram ajustados para se adequarem aos casos específicos, utilizando explicitamente as técnicas ajustadas como, por exemplo, para a redução de dimensionalidade com

UMAP e clusterização com HDBSCAN. Para cada adaptação diferente, os resultados obtidos foram analisados e comparados até que fossem encontrados os valores adequados dos parâmetros, o que fez com que o processo fosse o mais demorado em questões de implementação do que os outros.

É importante ressaltar que os resultados podem não ser consistentes entre as execuções devido à natureza estocástica da etapa de redução de dimensionalidade (UMAP). Para contornar tal comportamento é preciso que seja definido um valor na função do UMAP chamado `random_state`. Tal configuração não foi realizada no conjunto de dados do restaurante Guacamole, porém em todos os outros conjuntos.

A execução da análise de sentimentos, empregando um modelo disponível na plataforma do Hugging Face treinado com o BERT, foi utilizada somente para o conjunto de dados do restaurante Guacamole. Com os outros conjuntos, foi utilizado um modelo onde foi efetuado o treinamento através do BERTimbau, um modelo desenvolvido exclusivamente para a língua portuguesa, usando as avaliações da Olist, da Buscapé, da B2W Digital e UTLC. Vale ressaltar que, pela validação do modelo, o mesmo obteve uma acurácia de 78% e ambos retornam resultados com valores de “positivo”, “neutro” e “negativo”. Para avaliar o desempenho do modelo treinado, foi utilizada a função `classification_report`¹⁷ do scikit-learn¹⁸, uma biblioteca de aprendizado de máquina de código aberto para Python, que gera um relatório com métricas como precisão, sensibilidade, F1-score e suporte.

A extração de entidades nomeadas foi realizada utilizando um modelo disponível na plataforma do Hugging Face, treinado com o RoBERTa, em todos os conjuntos de dados como parte do método de análise. Esse modelo de linguagem é capaz de classificar entidades como localização, pessoa, organização, entre outras. Para garantir a eficácia, outros modelos de linguagem presentes no sistema, sendo em sua maioria treinados com o BERTimbau, foram inicialmente utilizados para validação manual nos conjuntos do Guacamole e da Olist, antes de optarmos por tal modelo final a ser empregado.

O método foi aplicado sequencialmente aos conjuntos de dados do Guacamole, da Olist, da Buscapé, da B2W Digital e dos hotéis brasileiros. Após a análise dos

¹⁷ Função `classification_report`. Disponível em: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.classification_report.html

¹⁸ Scikit-learn. Disponível em: <https://scikit-learn.org/>

resultados da Olist, foi identificada a necessidade de geração frases descritivas dos tópicos usando um modelo de linguagem avançado para facilitar ao usuário final o entendimento dos temas tratados no tópico. Para isso, desenvolvemos um comando (prompt) utilizando as 10 palavras mais importantes de cada tópico, junto com suas pontuações TF-IDF. Tal comando foi utilizado no ChatGPT¹⁹ e no Gemini²⁰ para a formação de três opções de frases, permitindo que seja possível avaliar qual delas melhor atendia às necessidades de interpretação. Essas frases foram essenciais para compreender o conteúdo e a relevância de cada tópico, destacando os termos-chave e sua importância contextual. Posteriormente, o comando foi refinado para o conjunto de dados B2W Digital, buscando melhorar ainda mais a qualidade das frases geradas para cada tópico identificado.

O primeiro comando elaborado foi: "Escreva, em uma única frase sucinta e descritiva, sobre o que se trata esse tópico obtido através de uma análise de tópicos, dando três frases alternativas." No entanto, ao analisar as respostas fornecidas pelo ChatGPT e Gemini, foi percebido que os modelos não consideraram o grau de importância de cada uma das palavras do tópico, juntamente com sua pontuação TF-IDF. Como resultado, o comando foi refinado para: "Escreva, em uma única frase sucinta e descritiva, sobre o que se trata esse tópico obtido através de uma análise de tópicos com BERTopic usando dados de [assunto do conjunto de dados], entregando três frases alternativas. Leve em consideração que as palavras que descrevem o tópico estão em ordem decrescente de relevância, do maior para o menor, no contexto do documento. Comece com 'O tópico trata sobre...'" Essa atualização tornou mais claro o objetivo desejado para as frases geradas dando mais importância às palavras com maior destaque nos tópicos obtidos.

Com o método finalmente consolidado e ajustado conforme necessário, o mesmo foi aplicado em todos os conjuntos de dados. Essa abordagem permitiu a visualização dos principais temas discutidos nas avaliações dos consumidores, bem como as relações entre seus sentimentos em relação a cada um desses temas, além das entidades mais frequentemente mencionadas. Com essa análise, é possível explorar os comentários associados a cada tipo de sentimento e identificar as entidades citadas com maior frequência nos comentários, oferecendo uma visão abrangente

¹⁹ ChatGPT. Disponível em: <https://chatgpt.com/>

²⁰ Gemini. Disponível em: <https://gemini.google.com/app>

das interações entre temas, sentimentos e entidades nos conjuntos de dados analisados.

Ao aplicar esse método de análise de textos em avaliações de consumidores, é possível obter informações cruciais para o aprimoramento da experiência do cliente e o desenvolvimento de estratégias direcionadas. Eis os principais insights que podem ser obtidos através das análises:



Figura 5 - Exemplos de insights obtidos pelo método. Fonte: A autora.

1. **Mapeamento de Assuntos e Interesses dos Clientes:** Identificação dos principais temas abordados nas avaliações, proporcionando insights sobre as preocupações e interesses dos clientes em relação aos produtos ou serviços.
2. **Percepção dos Sentimentos e Satisfação dos Clientes:** Categorização das opiniões dos consumidores como positivas, neutras ou negativas em relação aos temas identificados, oferecendo uma visão clara da satisfação geral e identificando áreas que necessitam de melhorias.
3. **Identificação de Entidades Relevantes:** Extração das entidades nomeadas mais mencionadas, como organizações, locais, produtos ou pessoas, revelando associações importantes feitas pelos clientes e partes cruciais de sua experiência.
4. **Julgamento dos Clientes pelos Assuntos:** Compreensão de como diferentes aspectos são vistos na percepção dos clientes, permitindo priorizar melhorias e

manter pontos fortes.

5. **Associações entre Assuntos e Entidades:** Exploração das relações entre temas discutidos nas avaliações e entidades nomeadas, revelando como organizações, locais ou pessoas estão associadas aos diferentes aspectos mencionados pelos clientes.
6. **Monitoramento da Reputação Online:** Utilização do monitoramento contínuo das avaliações para avaliar a reputação online da empresa e realizar ajustes estratégicos conforme necessário, visando manter ou melhorar a percepção dos clientes sobre a marca.

Essas análises combinadas oferecem uma visão holística da experiência do cliente e das percepções em relação aos produtos ou serviços. Com base nessas informações, os negócios podem desenvolver estratégias para aprimorar a satisfação do cliente, além de otimizar a oferta de produtos e serviços.

10. APLICAÇÃO DAS TÉCNICAS ESCOLHIDAS

O método foi aplicado em sete conjuntos de dados diferentes: nos dados do restaurante Guacamole, contidos na plataforma TripAdvisor; nas plataformas de comércio eletrônico Olist, Buscapé e B2W Digital; em hotéis brasileiros contidos no TripAdvisor e, de forma mais específica, no hotel Grand Hyatt de São Paulo, proveniente do conjunto de dados de hotéis brasileiros; e no smartphone da Motorola, proveniente do conjunto de dados da B2W Digital.

Em cada um dos sete casos, foi gerada uma nuvem de palavras para visualizar as palavras mais frequentes nos comentários. Em seguida, foram identificados os principais tópicos e, dos mais relevantes, selecionados para prosseguir com a aplicação do método. Para cada tópico, foram examinados tanto a quantidade de comentários associados quanto as 10 palavras mais importantes, acompanhadas de suas pontuações TF-IDF, para compreender o conteúdo e a relevância de cada tópico, destacando os termos-chave e sua importância dentro do contexto.

Com os tópicos selecionados, três opções de frases sucintas e informativas foram exibidas, cada uma descrevendo o tópico de forma única e clara com o auxílio das duas inteligências artificiais. Em seguida, a análise de sentimentos de cada tópico foi realizada, apresentando uma amostra dos comentários e as entidades nomeadas

de cada tópico. Com todas as informações coletadas, os comentários relacionados a cada tópico e sentimento foram exibidos, com a amostra composta pelos três primeiros comentários encontrados na tabela de dados.

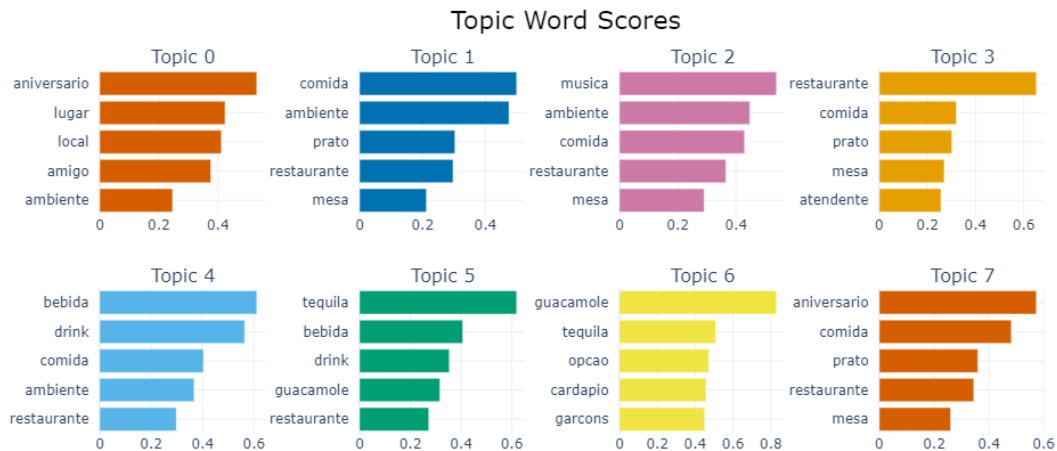
Por fim, após a aplicação do método, é apresentada uma análise técnica do método proposto na próxima seção, seguida pela avaliação do método por técnicos especializados na área de estudo. Ademais, ao final da exibição dos resultados da aplicação da metodologia no caso do restaurante, são demonstrados dois protótipos de apresentações, que são páginas web construídas com a utilização de HTML e CSS, de como os resultados seriam demonstrados a um usuário final.

10.1. RESTAURANTE GUACAMOLE

Para uma primeira aplicação do método, foi escolhido o conjunto de dados provenientes do restaurante Guacamole, obtidos por meio da plataforma do TripAdvisor em Dezembro de 2023, utilizando a técnica de web scraping. Escolhi esse conjunto de dados, composto por 2.872 avaliações, devido à sua quantidade relativamente reduzida, o que agiliza a análise e simplifica a compreensão dos resultados e da aplicação dos modelos de linguagem para as análises de mineração de texto.

Para que fosse possível ter uma demonstração visual das palavras mais mencionadas nos comentários das avaliações, foi gerada uma nuvem com tais palavras, onde é possível visualizar que palavras como “comida”, “atendimento”, “ambiente” e “restaurante” foram frequentemente escritas.

Prosseguindo com a modelagem de tópicos, foram identificados 8 tópicos, sendo que somente os 4 primeiros foram escolhidos para análise.



A seguir, é apresentada a aplicação completa do método para o tópico 0, incluindo as frases descritivas associadas. Também são mostrados os resultados da análise de sentimentos, com a quantidade de comentários classificados como “positivos”, “negativos” e “neutros”. Além disso, são listadas as entidades nomeadas encontradas e suas frequências. Por fim, são exibidos os três primeiros comentários relacionados ao tópico, categorizados por sentimento.

Tópico	Palavras do Tópico	Frequência
0	[('aniversario', 0.5316182), ('lugar', 0.42410007), ('local', 0.41171736), ('amigo', 0.3760357), ('ambiente', 0.2466469), ('preco', 0.22593665), ('noite', 0.22311413), ('decoracao', 0.21338886), ('atendimento', 0.20992842), ('atendente', 0.18922509)]	593

Modelo de Linguagem	Tópico	Frase
ChatGPT	0	O tópico trata sobre as experiências oferecidas pelo restaurante Guacamole para celebrações de aniversário, destacando o ambiente, local e atmosfera amigável.
ChatGPT	0	Este tópico aborda as opções do restaurante Guacamole para celebrações de aniversário, ressaltando o ambiente, local e clima acolhedor.
ChatGPT	0	Trata-se da experiência proporcionada pelo restaurante Guacamole em festas de aniversário, com ênfase no ambiente, localização e atmosfera amigável.
Gemini	0	O tópico trata sobre as comemorações de aniversário no restaurante Guacamole, com foco em ambiente aconchegante, preços acessíveis, boa localização e atendimento atencioso para tornar a noite especial.
Gemini	0	O tópico aborda a experiência de aniversários no Guacamole, incluindo

		ambiente agradável, decoração temática, opções para grupos de amigos e qualidade do atendimento para aniversariantes e convidados.
Gemini	0	O tópico foca em celebrar aniversários no Guacamole, mencionando ambiente convidativo, preços justos, localização, facilidade para reunir amigos e equipe atenciosa para proporcionar uma experiência memorável.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
0	548	21	24

Tópico	Entidades frequentes	Frequência
0	mexicana, mexico, rio, mexicano, garcom, gabriel, andre, janeiro, alisson, lucas	20, 12, 11, 9, 8, 8, 6, 4, 4, 4

Tópico	Comentário	Sentimento
0	Amei o atendimento da Kaylane e Victor! Os melhoress, comemos muito bem e foram muitos simpáticos. Obrigada pela ajuda	positivo
0	Lugar e pessoas muito gentis e educadas. Fomos atendidos por um garçom incrível, Mateus. Vamos voltar sem dúvidas!	positivo
0	Excelente atendimento, amei a experiência! Rodízio com bastante variedade e comida deliciosa, só tenho a agradecer	positivo
0	Este local é demasiado animado para curtir a noite em	neutro

	família nessa cidade maravilhosa que é o Rio de Janeiro	
0	Adorei o clima mexicano!! Atendimento do Marlon foi o melhor de todos! Fizemos bagunça e mesmo assim pessoal atendeu bem.	neutro
0	Recomendo a todos que venham aqui porque a animação é garantida e vai dar tudo certo. Muito bom, Hermanos.	neutro
0	Falta de organização, + de 1 hora de rodizio e só chegou duas tortinhas e uma costela, lindao esse rodízio	negativo
0	Após chegar com minha família fomos tratados com descaso e de forma irônica por termos uma forma peculiar no sotaque,infelizmente não foi leve,soou grosseiro!!!	negativo
0	Estive com um grupo de amigos (13 para ser exato),para aniversário de uma das. Fomos na recepção, e a atendente informou que não faziam reservas aos fins de semana (era sábado), e que era difícil alocar um número grande, mas seríamos chamados. Vimos várias pessoas...	negativo

A seguir, é apresentada a aplicação completa do método para o tópico 1, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
1	[('comida', 0.4991731), ('ambiente', 0.47494364), ('prato', 0.30317855), ('restaurante', 0.29707727), ('mesa', 0.21258432), ('decoracao', 0.15873152), ('atendimento', 0.15360776), ('local', 0.1474943), ('lugar', 0.13329042),	570

	('preco', 0.12596495)]	
--	------------------------	--

Modelo de Linguagem	Tópico	Frase
ChatGPT	1	O tópico trata sobre a experiência gastronômica e o ambiente oferecido pelo restaurante Guacamole, destacando a qualidade da comida e a atmosfera do local.
ChatGPT	1	Este tópico aborda a culinária e o ambiente do restaurante Guacamole, enfatizando a qualidade da comida e a atmosfera.
ChatGPT	1	Trata-se da experiência gastronômica e o ambiente do restaurante Guacamole, com ênfase na comida de alta qualidade e na atmosfera do local.
Gemini	1	O tópico trata sobre a experiência gastronômica no restaurante Guacamole, com foco em comida saborosa, ambiente agradável, variedade de pratos, atendimento atencioso e preços justos.
Gemini	1	O tópico aborda as qualidades do restaurante Guacamole, incluindo comida deliciosa, ambiente acolhedor, cardápio diversificado, equipe atenciosa e custo-benefício, tornando-o uma ótima opção para uma boa refeição.
Gemini	1	O tópico foca em desfrutar de uma boa comida no Guacamole, mencionando pratos saborosos, ambiente convidativo, atendimento de qualidade, variedade de opções no menu e preços acessíveis para

		agradar a todos os paladares.
--	--	-------------------------------

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
1	525	25	20

Tópico	Entidades frequentes	Frequência
1	mexicana, mexicano, mexico, garcom, matheus, rio, gabriel, mari, andre, janeiro	77, 15, 7, 7, 6, 5, 4, 3, 3, 3

Tópico	Comentário	Sentimento
1	Atendimento maravilhoso! Matheus foi muito atencioso e colocou nossa playlist. Todos os atendentes super simpáticos	positivo
1	Gostei muito dos pratos e do atendimento. Muitos solícitos. A comida estava saborosa. Com certeza volto mais vezes!	positivo
1	Eu amei, é a melhor cozinha mexicana que eu já comi. E o espaço é muito animado, imersivo! A decoração da casa é linda e o atendimento excelente sempre.	positivo
1	Comida boa, rodízio completo porém muiiiita demora entre um prato e outro, o atendimento é bom, pessoal simpático.	neutro
1	O local é bem bonito, bom atendimento, mas achei pouca variedade no rodízio. Restringindo alguns pratos. O sabor poderia ser melhor tb. Apesar de ser comida mexicana quase não tem pimenta.	neutro
1	Seria legal ter Taco recheado e opções com camarão e não vem.os.pedidos feitos e só os que vem a mesa.	neutro

1	Não tive uma boa experiência. Atendimento deixou a desejar, tivemos que pedir 3x um determinado prato. Rodízio sempre servindo os mesmos pratos. Não fomos avisados que possuíam um mesa onde também podíamos nos servir, nos demos conta porque a mesa ao lado se serviu...	negativo
1	A comida não é ruim, mas os garçons são tentos e o caixa é péssimo, lento e confuso . Mesmo com a casa relativamente vazia, tudo demorou muito	negativo
1	Sábado dia 25/05/19 fui ao guacamole com minha esposa, tinha uma fila na porta, fui pedir para colocar nosso nome na lista de espera e o organizador da fila disse que não tinha mesa para 2 pessoas, somente 3 ou mais. Falei que aguardaria liberar...	negativo

A seguir, é apresentada a aplicação completa do método para o tópico 2, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
2	[('musica', 0.5390509), ('ambiente', 0.44739813), ('comida', 0.4295042), ('restaurante', 0.36548102), ('mesa', 0.29080635), ('prato', 0.25236893), ('decoracao', 0.2010031), ('lugar', 0.18700764), ('local', 0.18374026), ('bebida', 0.1709653)]	519

Modelo de Linguagem	Tópico	Frase
ChatGPT	2	O tópico trata sobre a experiência sensorial no restaurante Guacamole, envolvendo música, ambiente agradável e comida saborosa.

ChatGPT	2	Esse tópico aborda a atmosfera musical e o ambiente acolhedor do restaurante Guacamole, destacando também a qualidade da comida.
ChatGPT	2	Trata-se da experiência completa no restaurante Guacamole, incluindo música ambiente, atmosfera acolhedora e deliciosos pratos.
Gemini	2	O tópico trata sobre experiência musical e gastronômica no restaurante Guacamole, com foco em música ambiente, ambiente agradável, comida saborosa e variedade de pratos para um momento completo.
Gemini	2	O tópico aborda a combinação de música, comida e ambiente no Guacamole, incluindo música ao vivo, atmosfera descontraída, cardápio diversificado e espaço aconchegante para uma noite memorável.
Gemini	2	O tópico foca em curtir música e boa comida no Guacamole, mencionando trilha sonora agradável, ambiente convidativo, pratos deliciosos e opções para todos os gostos, proporcionando uma experiência única aos clientes.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
2	497	24	8

Tópico	Entidades frequentes	Frequência
---------------	-----------------------------	-------------------

2	mexicana, mexicano, mexico, garcom, rio, matheus, janeiro, brasileira, andre, lucas	87, 30, 16, 9, 7, 5, 5, 3, 3, 3
---	---	---------------------------------

Tópico	Comentário	Sentimento
2	Ótimo ambiente e decoração, funcionários prestativos, comida boa. Música agradável. Divertidos mariachis.	positivo
2	Ambiente agradável, atendimento nota 10, cozinha maravilhosa, chopp gelado, drinks maravilhosos, música ótima, tudo com tempero mexicano	positivo
2	Ótimo ambiente, comida e atendimento! Tudo muito animado, músicas e artistas temáticos com bastante energia!	positivo
2	O rodízio é interessante, mas nada fora do normal, faltou um pouco de pimenta. O ambiente é legal, mas a música junina não foi muito agradável, apesar de entender que é a temática do mês nos restaurantes.	neutro
2	Os mariachis cantaram Rebelde pra nossa mesa e a gente se sentiu a própria Anahí! Eles também cantaram Cielito Lindo, galera mexicana. Artistas muito talentosos e desenrolados. Atendimento top! A comida tava excelente, vou passar os proximos dias lembrando das quesadillas e do taco	neutro
2	Fui no aniversário de uma amiga e nos divertimos muito, com os Mariachis, a comida, drinks e ao final virou uma discoteca. Caso não goste de música, melhor verificar os dias no ato da reserva ou entrada.	neutro
2	Atendimento ruim. A comida demorou demais e quando chegou, o garçom se enganou de mesa e tivemos que aguardar até que a confusão fosse resolvida. O que salvou foi a comida excelente e os três músicos do	negativo

	restaurante.	
2	Péssimo atendimento, pedidos muito demorados, os garçons estavam sem vontade e quando vinham atender parecia que estavam fazendo favor.	negativo
2	O ambiente, comida e bebida são muito bons, porém o atendimento é horrível, garçons, atendentes e segurança todos mal-educados e despreparados, da primeira vez que tentei ir, apesar de estar com reserva tivemos que esperar tanto na fila que acabamos desistindo...	negativo

A seguir, é apresentada a aplicação completa do método para o tópico 3, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
3	[('restaurante', 0.6527969), ('comida', 0.32020435), ('prato', 0.30163592), ('mesa', 0.26993367), ('atendente', 0.25612783), ('decoracao', 0.2269316), ('local', 0.14462337), ('opcao', 0.14196447), ('aniversario', 0.13936025), ('ambiente', 0.13637745)]	350

Modelo de Linguagem	Tópico	Frase
ChatGPT	3	O tópico trata sobre a experiência geral no restaurante Guacamole, abrangendo o ambiente, comida, variedade de pratos e atendimento.
ChatGPT	3	Este tópico discute o restaurante Guacamole, incluindo aspectos como comida, variedade de pratos, ambiente e atendimento.

ChatGPT	3	Trata-se da análise do restaurante Guacamole, destacando sua comida, ambiente, opções de pratos e o atendimento oferecido.
Gemini	3	O tópico trata sobre a experiência completa no restaurante Guacamole, com foco em comida deliciosa, ambiente agradável, atendimento atencioso e diversas opções para atender às necessidades dos clientes.
Gemini	3	O tópico aborda as qualidades do restaurante Guacamole, incluindo pratos saborosos, ambiente acolhedor, equipe atenciosa, variedade de opções no menu e um local propício para diversas ocasiões, como aniversários.
Gemini	3	O tópico foca em desfrutar de uma boa refeição no Guacamole, mencionando comida de qualidade, ambiente convidativo, atendimento de qualidade e diversas opções para agradar a todos os paladares, tornando-o uma ótima escolha para um momento especial.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
3	304	23	23

Tópico	Entidades frequentes	Frequência
3	mexicana, mexicano, rio, janeiro, mexico, garcom, brasil, olegario, alisson, rio	49, 32, 15, 9, 9, 6, 5, 4, 3, 3

Tópico	Comentário	Sentimento
3	Restaurante agradável, fomos muito bem atendidos. Os pedidos vieram rápido e a comida estava bem gostosa.	positivo
3	Achei mais do que um restaurante e sim uma experiência absolutamente fora do padrão! Amei o di a comida a bebida e o espaço	positivo
3	Bom restaurante, ambiente com decoração e canções típicas, e bom atendimento. Preço justo e tudo a la carte. A rua onde fica é quase impossível achar vaga para estacionar.	positivo
3	Nos arrependemos de optar pelo rodízio, pois além da demora em servir, não tinham tantas opções. A la carte é uma boa, pois há mais opções, no rodízio eles restringem. No geral um restaurante gostoso, ambiente agradável.	neutro
3	Restaurante muito bonito, garçons simpáticos, em especial a funcionaria Maria. Voltaria outras vezes. Nota 10!!!	neutro
3	Restaurante é bem localizado e decorado. Atendimento bom. Mas obque pedimos, quesadillas, estavam muito oleosas. Minha filha pediu um nacho e estava melhor.	neutro
3	Mesmo num dia chuvoso, restaurante bastante quente e abafado, ar condicionado não funciona ou não dá vazão! Funcionário do caixa me cobrou valor a mais com arrogância, se eu não noto o erro iria ficar por isso mesmo! Bebidas fracas, mais pareciam refresco! Não gostei!	negativo
3	Comida, ao menos no dia em que fui, muito aquém da fama do restaurante. A ambientação era melhor que o tempero. Guacamole, que dá nome ao restaurante,	negativo

	fraca. Espero retornar e mudar de opinião.	
3	Se quiser diversão não tenha dúvida de onde ir. Tequileira durante toda a noite. No final da noite o restaurante torna-se quase uma boate, com música animada, personagens... impossível ficar parado. Apenas a comida deixou a desejar.	negativo

A seguir, apresentam-se os protótipos de apresentações em formato de páginas web que exibem os resultados obtidos para os casos dos tópicos de número 0 e 1, respectivamente.



Figura 8 - Exibição das informações gerais do tópico 0 contidas no protótipo de apresentação. Fonte: A autora.

Sentimentos dos Comentários	
Sentimento	Quantidade
Positivos	548
Neutros	21
Negativos	24

Figura 9 - Exibição das informações da análise de sentimentos do tópico 0 contidas no protótipo de apresentação. Fonte: A autora.

Entidades Nomeadas Frequentes	
Entidade	Frequência
mexicana	20
mexico	12
rio	11
mexicano	9
garcom	8
gabriel	8
andre	6
janeiro	4
alisson	4
lucas	4

Figura 10 - Exibição das informações da extração de entidades nomeadas do tópico 0 contidas no protótipo de apresentação. Fonte: A autora.

Exemplos de Comentários	
Sentimento: Positivo	Amei o atendimento da Kaylane e Victor! Os melhoress, comemos muito bem e foram muitos simpáticos. Obrigada pela ajuda
Sentimento: Positivo	Lugar e pessoas muito gentis e educadas. Fomos atendidos por um garçom incrível, Mateus. Vamos voltar sem dúvidas!
Sentimento: Positivo	Excelente atendimento, amei a experiência! Rodízio com bastante variedade e comida deliciosa, só tenho a agradecer

Figura 11 - Exibição dos três primeiros comentários positivos do tópico 0 contidos no protótipo de apresentação. Fonte: A autora.

Sentimento: Neutro	Este local é demasiado animado para curtir a noite em família nessa cidade maravilhosa que é o Rio de Janeiro
Sentimento: Neutro	Adorei o clima mexicano!! Atendimento do Marlon foi o melhor de todos! Fizemos bagunça e mesmo assim pessoal atendeu bem.
Sentimento: Neutro	Recomendo a todos que venham aqui porque a animação é garantida e vai dar tudo certo. Muito bom, Hermanos.

Figura 12 - Exibição dos três primeiros comentários neutros do tópico 0 contidos no protótipo de apresentação. Fonte: A autora.

Sentimento: Negativo	Falta de organização, + de 1 hora de rodízio e só chegou duas tortinhas e uma costela, lindao esse rodízio
Sentimento: Negativo	Após chegar com minha família fomos tratados com descaso e de forma irônica por termos uma forma peculiar no sotaque, infelizmente não foi leve, soou grosseiro!!!
Sentimento: Negativo	Estive com um grupo de amigos (13 para ser exato), para aniversário de uma das. Fomos na recepção, e a atendente informou que não faziam reservas aos fins de semana (era sábado), e que era difícil alocar um número grande, mas seríamos chamados. Vimos várias pessoas...

Figura 13 - Exibição dos três primeiros comentários negativos do tópico 0 contidos no protótipo de apresentação. Fonte: A autora.

Por fim, o protótipo de página relacionado ao tópico 1.

Análise de Comentários - Restaurante Guacamole

Tópico 1: Experiência Gastronômica

Total de Comentários no Tópico: 570

Descrição do Tópico: O tópico trata sobre a experiência gastronômica no restaurante Guacamole, com foco em comida saborosa, ambiente agradável, variedade de pratos, atendimento atencioso e preços justos.

Nuvem de Palavras do Tópico



Figura 14 - Exibição das informações gerais do tópico 1 contidas no protótipo de apresentação. Fonte: A autora.

Sentimentos dos Comentários

Sentimento	Quantidade
Positivos	525
Neutros	25
Negativos	20

Figura 15 - Exibição das informações da análise de sentimentos do tópico 1 contidas no protótipo de apresentação. Fonte: A autora.

Entidades Nomeadas Frequentes

Entidade	Frequência
mexicana	77
mexicano	15
mexico	7
garcom	7
matheus	6
rio	5
gabriel	4
mari	3
andre	3
janeiro	3

Figura 16 - Exibição das informações da extração de entidades nomeadas do tópico 1 contidas no protótipo de apresentação. Fonte: A autora.

Exemplos de Comentários

Sentimento: Positivo

Atendimento maravilhoso! Matheus foi muito atencioso e colocou nossa playlist. Todos os atendentes super simpáticos

Sentimento: Positivo

Gostei muito dos pratos e do atendimento. Muitos solícitos. A comida estava saborosa. Com certeza volto mais vezes!

Sentimento: Positivo

Eu amei, é a melhor cozinha mexicana que eu já comi. E o espaço é muito animado, imersivo! A decoração da casa é linda e o atendimento excelente sempre.

Figura 17 - Exibição dos três primeiros comentários positivos do tópico 1 contidos no protótipo de apresentação. Fonte: A autora.

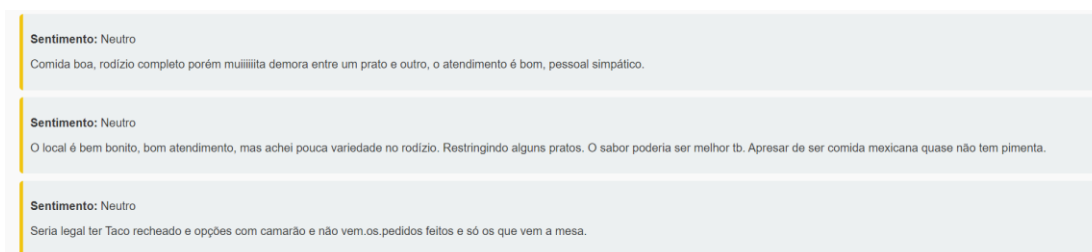


Figura 18 - Exibição dos três primeiros comentários neutros do tópico 1 contidos no protótipo de apresentação. Fonte: A autora.

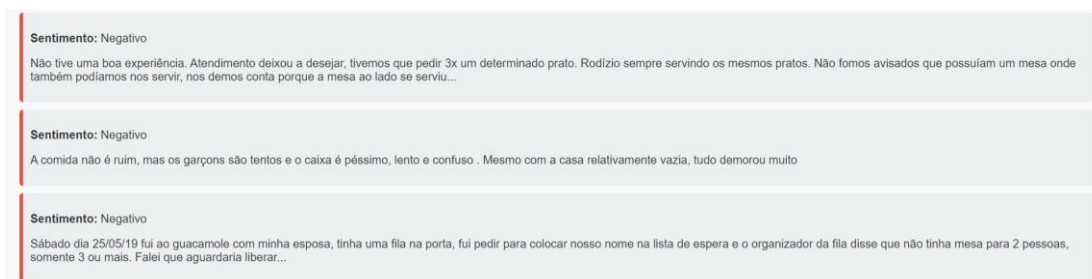


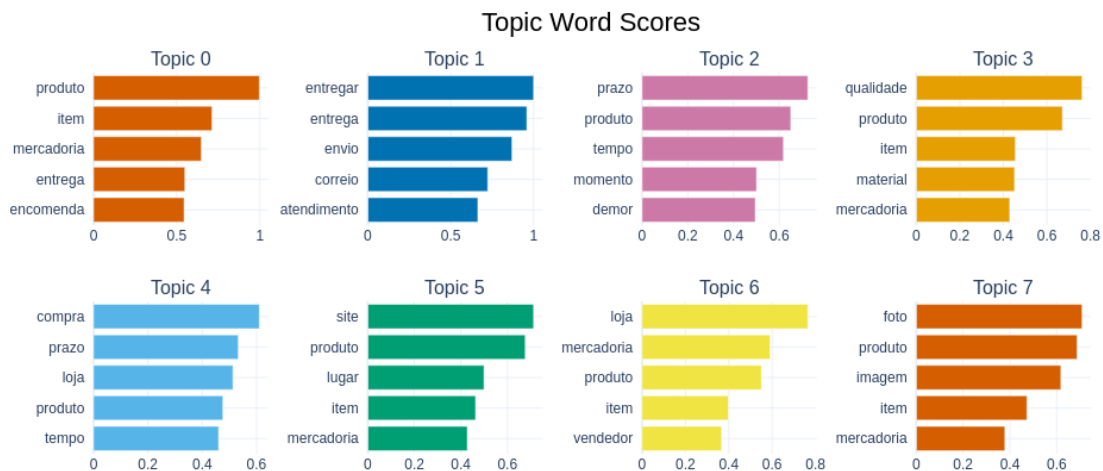
Figura 19 - Exibição dos três primeiros comentários negativos do tópico 1 contidos no protótipo de apresentação. Fonte: A autora.

10.2. OLIST

O método foi aplicado ao conjunto de dados da plataforma de comércio eletrônico Olist, que foi disponibilizado pela própria empresa brasileira na plataforma do Kaggle. Esse conjunto abrange aproximadamente 100 mil avaliações coletadas entre os anos de 2016 e 2018. Durante a análise desse conjunto de dados, foi observado que várias avaliações não continham comentários, consistindo apenas na atribuição de um número de estrelas (variando de 1 a 5). Portanto, foram consideradas apenas as avaliações que continham comentários válidos, resultando em um total de 37.564 comentários.

Para que fosse possível ter uma demonstração visual das palavras mais mencionadas nos comentários das avaliações foi gerada uma nuvem com tais palavras, onde é possível visualizar que palavras como “produto”, “prazo”, “qualidade” e “compra” foram frequentemente escritas.

Prosseguindo com a modelagem de tópicos, foram identificados 48 tópicos, sendo que somente os 4 primeiros foram escolhidos para análise.



A seguir, é apresentada a aplicação completa do método para o tópico 0, incluindo as frases descritivas associadas. Também são mostrados os resultados da análise de sentimentos, com a quantidade de comentários classificados como “positivos”, “negativos” e “neutros”. Além disso, são listadas as entidades nomeadas encontradas e suas frequências. Por fim, são exibidos os três primeiros comentários relacionados ao tópico, categorizados por sentimento.

Tópico	Palavras do Tópico	Frequência
0	[('produto', 0.99698925), ('item', 0.7123114), ('mercadoria', 0.64804196), ('entrega', 0.548958), ('encomenda', 0.5444926), ('marca', 0.54118776), ('entregar', 0.52260983), ('aparelho', 0.5213896), ('fornecedor', 0.51684713), ('empresa', 0.5129562)]	5482

Modelo de Linguagem	Tópico	Frase
ChatGPT	0	O tópico trata sobre a análise de produtos e itens comercializados na plataforma Olist, com foco em aspectos como mercadoria, entrega, marca e fornecedores.
ChatGPT	0	Este tópico aborda a variedade de produtos e itens disponíveis no Olist, destacando aspectos como entrega, marca e fornecedores.
ChatGPT	0	Trata-se da análise dos produtos e mercadorias oferecidos pelo Olist, com ênfase na entrega, marcas e diversidade de fornecedores.
Gemini	0	O tópico trata sobre a gestão de produtos na plataforma Olist, com foco em cadastro de itens, entrega de mercadorias, marcas dos produtos e fornecedores da plataforma.
Gemini	0	O tópico aborda os desafios e oportunidades da logística no e-commerce, utilizando a Olist como exemplo, desde o cadastro de produtos até a entrega final ao

		cliente.
Gemini	0	O tópico foca em otimizar as operações de venda na Olist, mencionando o cadastro eficiente de produtos, processo de entrega ágil, variedade de marcas e relacionamento com fornecedores.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
0	2720	201	2561

Tópico	Entidades frequentes	Frequência
0	correios, targaryen, china, lannister, asus, zenfone, samsung, sao, baratheon, invicta	13, 8, 5, 5, 4, 4, 4, 4, 4, 3

Tópico	Comentário	Sentimento
0	Produto corresponde ao descrito, é ótimo!	positivo
0	Excelente produto, chegou antes do que eu imaginava!	positivo
0	Um ótimo produto!	positivo
0	O produto é bom, porém imaginava que fosse maior..	neutro
0	Nada mais a avaliar, pois ainda estou testando o produto!	neutro
0	O produto deveria vir com parafusos.	neutro
0	O produto veio sem manual ensinando como montar.	negativo
0	Era para receber o produto até 19/3 e age hoje (23/3) não recebi e não sei o motivo.	negativo
0	O carrinho não foi entregue só a boneca. Qdo bem???	negativo

A seguir, é apresentada a aplicação completa do método para o tópico 1, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
1	[('entregar', 0.99999994), ('entrega', 0.9592142), ('envio', 0.86956406), ('correio', 0.7232951), ('atendimento', 0.6638352), ('pedir', 0.65225315), ('endereço', 0.64826906), ('pedido', 0.63752705), ('serviço', 0.6287203), ('condicao', 0.60464996)]	5430

Modelo de Linguagem	Tópico	Frase
ChatGPT	1	O tópico trata sobre a logística de entrega e envio na plataforma Olist, abordando aspectos como atendimento ao cliente, pedidos e serviços relacionados.
ChatGPT	1	Este tópico discute a eficiência e operação de entrega na plataforma Olist, incluindo serviços de atendimento, pedidos e condições de envio.
ChatGPT	1	Trata-se da análise detalhada do processo de entrega e envio no Olist, destacando o atendimento ao cliente, pedidos e serviços associados.
Gemini	1	O tópico trata sobre a eficiência na entrega de pedidos na Olist, com foco em prazos de entrega, logística de envio e serviços de correio.
Gemini	1	O tópico aborda os desafios e soluções da entrega de produtos no e-commerce, utilizando a Olist como exemplo, desde o

		pedido até a entrega final ao cliente.
Gemini	1	O tópico foca em garantir uma ótima experiência de entrega aos clientes da Olist, mencionando prazos ágeis, rastreamento de pedidos, atendimento de qualidade e condições flexíveis.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
1	4012	299	1119

Tópico	Entidades frequentes	Frequência
1	ótima, confiavel, brasil, baratheon, linda, cuba, joia, razoavel, valeu, correios	28, 8, 6, 5, 5, 4, 3, 3, 3, 3

Tópico	Comentário	Sentimento
1	Recomendo	positivo
1	tudo conforme combinado	positivo
1	Tudo ok!	positivo
1	Estou no aguardo	neutro
1	No quesito entrega vocês estão de parabéns. Chegou muito antes do previsto. Pena que o tablet trava muito e o sistema usa muito do espaço interno que já não é lá essas coisas.	neutro
1	Achei um pouco fragil	neutro
1	Comprei dois arranhadores e somente um foi entregue	negativo
1	Não recebi o espremedor	negativo
1	Eu não recebi o peoduto	negativo

A seguir, é apresentada a aplicação completa do método para o tópico 2, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
2	[('prazo', 0.72487533), ('produto', 0.65022784), ('tempo', 0.61835593), ('momento', 0.50057465), ('demor', 0.4949626), ('atraso', 0.4807905), ('hora', 0.46097136), ('item', 0.42442542), ('preco', 0.38504636), ('rapidez', 0.38150612)]	4963

Modelo de Linguagem	Tópico	Frase
ChatGPT	2	O tópico trata sobre a gestão de prazos e tempo de entrega na plataforma Olist, abordando questões como eficiência operacional, experiência do cliente e gestão de estoque.
ChatGPT	2	Este tópico analisa a gestão de prazos e tempo de entrega na plataforma Olist, destacando a importância da eficiência operacional e da experiência do cliente.
ChatGPT	2	Trata-se da análise dos prazos de entrega e gestão do tempo na Olist, abordando aspectos como eficiência operacional e satisfação do cliente.
Gemini	2	O tópico trata sobre a gestão de prazos de entrega na Olist, com foco em otimizar o tempo de entrega dos produtos aos clientes.
Gemini	2	O tópico aborda os desafios de cumprir

		prazos de entrega no e-commerce, utilizando a Olist como exemplo, desde a logística até a comunicação com o cliente.
Gemini	2	O tópico foca em garantir entregas rápidas e eficientes na Olist, mencionando prazos curtos, acompanhamento do pedido e comunicação transparente com o cliente.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
2	4153	89	721

Tópico	Entidades frequentes	Frequência
2	baratheon, targaryen, correios, rio, barathe, ios, sao, paulo, baratheonr, belo	9, 5, 4, 2, 2, 2, 1, 1, 1, 1

Tópico	Comentário	Sentimento
2	Perfeito....chegou antes do prazo.....	positivo
2	Chegou bem antes do tempo e em otimas condicoes	positivo
2	Foi entregue antes do prazo.	positivo
2	achei o prazo da entrega um pouco demorado...contudo, suponho que a greve dos caminhoneiros deva ter afetado de alguma forma...	neutro
2	O produto foi o correto, chegou no prazo, entretanto houve muita confusão no rastreamento do produto, passando muitos dias para atualizar uma localização que fisicamente dura 90 minutos.	neutro
2	Apesar de estar no prazo de entrega o produto demorou muito para chegar, fiz o pedido do mesmo produto em	neutro

	um concorrente do stark que entregou bem antes.	
2	O produto foi entregue em tempo mas ao tirar do pacote e abrir quebrou.	negativo
2	A entrega foi demorada mas dentro do prazo. Mas a mochila não é o eu esperava, por ser de marca eu achei muito simples , não gostei nem um pouco do produto.	negativo
2	O produto não chegou e nem tem um código rastreador para que eu possa verificar o status da entrega, o prazo já passou e nada.	negativo

A seguir, é apresentada a aplicação completa do método para o tópico 3, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
3	[('qualidade', 0.76154006), ('produto', 0.67129004), ('item', 0.45518458), ('material', 0.4513951), ('mercadoria', 0.4293698), ('respeito', 0.3742519), ('marca', 0.35110795), ('super', 0.32365304), ('satisfacao', 0.31821865), ('fornecedor', 0.3143722)]	1841

Modelo de Linguagem	Tópico	Frase
ChatGPT	3	O tópico trata sobre a avaliação da qualidade dos produtos na plataforma Olist, considerando aspectos como materiais, marcas e satisfação do cliente.
ChatGPT	3	Este tópico analisa a qualidade dos produtos disponíveis na plataforma Olist, com foco em materiais, marcas e satisfação dos clientes.
ChatGPT	3	Trata-se da análise da qualidade dos

		produtos oferecidos na Olist, abordando características como materiais, marcas e níveis de satisfação dos consumidores.
Gemini	3	O tópico trata sobre a qualidade dos produtos na Olist, com foco em materiais utilizados, marcas confiáveis e alto nível de satisfação dos clientes.
Gemini	3	O tópico aborda a importância de produtos de qualidade no e-commerce, utilizando a Olist como exemplo, desde a seleção de fornecedores até a experiência do cliente final.
Gemini	3	O tópico foca em garantir a qualidade e confiabilidade dos produtos vendidos na Olist, mencionando materiais duráveis, marcas renomadas, fornecedores confiáveis e alto nível de satisfação entre os consumidores.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
3	1523	65	253

Tópico	Entidades frequentes	Frequência
3	baratheon, lannister, samsung, americana, targaryen, logan, nnister, luciane, apiai, eua	7, 3, 3, 2, 2, 1, 1, 1, 1, 1

Tópico	Comentário	Sentimento
3	Produto chegou antes do esperado e é de boa qualidade. Super gostei.	positivo

3	Tudo certo com a entrega e o produto é de boa qualidade	positivo
3	Recebiendo o produto no prazo produtos de Qualidade eu recom	positivo
3	O produto até foi entregue dentro do prazo mas, demorou muito, e recebi o produto com a cor errada e com o visor arranhado. O produto (fone) de boa qualidade.	neutro
3	O produto foi entregue corretamente, e até antes do prazo, porém, a sua qualidade é bem inferior ao esperado.	neutro
3	Achei que a largura e a qualidade fosse superior	neutro
3	PRAZO DE ENTREGA OK QUALIDADE DO PRODUTO OK FALTOU APENAS A FICHA DO "DEALER" A QUAL CONSTA NA CAIXA E NÃO CHEGOU.	negativo
3	Quero aqui dizer que sou cliente de muito tempo da stark e compro muito...Mais nunca tinha comprado um produto de tão péssima qualidade como esse...Quero aqui dizer que o produto vai pra lixo...	negativo
3	Achei de péssima qualidade	negativo

10.3. BUSCAPÉ

O método foi aplicado ao conjunto de dados da plataforma de comércio eletrônico Buscapé, que foi disponibilizado pelo escritório de pesquisa da USP através do projeto Opinando, em seu site. Esse conjunto abrange aproximadamente 84 mil avaliações coletadas no ano de 2013. Durante a análise desse conjunto de dados, foi observado que várias avaliações não continham comentários, consistindo apenas na atribuição de um número de estrelas (variando de 1 a 5). Portanto, foram consideradas apenas as avaliações que continham comentários válidos, resultando em um total de 79.025 comentários.

incluindo as frases descritivas associadas. Também são mostrados os resultados da análise de sentimentos, com a quantidade de comentários classificados como “positivos”, “negativos” e “neutros”. Além disso, são listadas as entidades nomeadas encontradas e suas frequências. Por fim, são exibidos os três primeiros comentários relacionados ao tópico, categorizados por sentimento.

Tópico	Palavras do Tópico	Frequência
0	[('qualidade', 0.5860331), ('produto', 0.53030264), ('custo', 0.29981625), ('desempenho', 0.26627156), ('camera', 0.2350088), ('preço', 0.2282173), ('som', 0.2275001), ('equipamento', 0.2259602), ('smartphone', 0.22288465), ('aparelho', 0.21339001)]	40452

Modelo de Linguagem	Tópico	Frase
ChatGPT	0	O tópico trata sobre a avaliação de produtos em termos de qualidade, design e custo na plataforma de comércio eletrônico Buscapé.
ChatGPT	0	Este tópico aborda a análise de produtos em relação à qualidade, design e custo na plataforma Buscapé.
ChatGPT	0	Trata-se da avaliação de produtos em termos de qualidade, design e custo na plataforma de comércio online Buscapé
Gemini	0	O tópico trata sobre a avaliação de produtos no Buscapé, com foco em qualidade, design, custo, desempenho e expectativas dos consumidores.
Gemini	0	O tópico aborda a comparação de produtos no Buscapé, incluindo qualidade, design, preço, desempenho e expectativas

		dos usuários, ajudando na decisão de compra.
Gemini	0	O tópico foca na experiência do consumidor com produtos no Buscapé, mencionando qualidade, design, custo-benefício, desempenho e atendimento às expectativas, visando a satisfação dos clientes.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
0	35311	2472	2669

Tópico	Entidades frequentes	Frequência
0	samsung, sony, android, nokia, brasil, iphone, windows, galaxy, son, apple	951, 793, 587, 442, 329, 326, 294, 281, 240, 216

Tópico	Comentário	Sentimento
0	Estou muito satisfeito, o visor é melhor do que eu imaginava, boas imagens, desing ultra fino. Pelo preço é um exelente aparelho. O que gostei: Desing exelente, display, custo beneficio. O que não gostei: Não tem como adicionar mais papeis de parede, bateria dura pouco.	positivo
0	"muito boa O que gostei: preco O que não gostei: poderia ser 42	positivo
0	Produto de ótima qualidade em todos os quesito!	positivo
0	É um bom produto, só precisa melhorar o design.	neutro
0	ele e bom pq serve só pra ligar a câmera esós as vezes	neutro

	que ela e boa mais na maioria das vezes O que gostei: E pq ele tem entrda pra chip O que não gostei: a câmera o touch market	
0	O viva-voz é bom, só as vezes da falha. O microfone do viva-voz é ruim, tem que ficar com a boca bem perto para ser ouvido pela outra pessoa. Já o auto-falante é muito bom!!! Já sem o viva-voz ativo o auto-falante fica muito baixo. E a bateria dura muito pouco, mais ou menos uns 70% menos que o anunciado!!! O que gostei: a qualidade do viva-voz na maioria das vezes é boa O que não gostei: a bateria durta muito MENOS que o anunciado	neutro
0	nao comparia novamente essa marca, pois paguei caro e me decepcionei com a durabilidade O que gostei: se vc for ver e analisar, parece uma ótima tv, porem comprei uma lcd 42 full hd, conversor, e que durou apenas 1 ano e meio, alem da assistencia técnica pessima, o conserto nao fica abaixo dos 700,00, só entrar no site "reclame aqui" e ver como essa marca da problema em televisores..cuidado fiquem alerta O que não gostei: a durabilidade	negativo
0	Faz mto barulho. O que gostei: Pequeno. O que não gostei: Faz mto ruído.	negativo
0	e muito lento pra conectar e ainda travou o que e que eu fasso pra colocar a vessao original.	negativo

A seguir, é apresentada a aplicação completa do método para o tópico 1, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
1	[('forno', 0.48017842), ('cozinha', 0.34369832), ('queimador', 0.33650964),	3499

	('grill', 0.24332425), ('prato', 0.20231475), ('fogao', 0.1818408), ('gas', 0.1672973), ('bolo', 0.13586965), ('aquecedor', 0.12980962), ('produto', 0.12527646)]	
--	---	--

Modelo de Linguagem	Tópico	Frase
ChatGPT	1	O tópico trata sobre análises de produtos relacionados à cozinha, como fornos, fogões e queimadores, disponíveis na plataforma de comércio eletrônico Buscapé.
ChatGPT	1	Este tópico aborda análises de produtos de cozinha, incluindo fornos, fogões e queimadores, na plataforma Buscapé.
ChatGPT	1	Trata-se da avaliação de produtos para cozinha, como fornos, fogões e queimadores, na plataforma de comércio online Buscapé.
Gemini	1	O tópico trata sobre fornos e eletrodomésticos para cozinha no Buscapé, com foco em modelos de forno, funções, queimadores, grill e praticidade no preparo de alimentos.
Gemini	1	O tópico aborda a escolha do forno ideal no Buscapé, incluindo tipos de fornos, funções, marcas, preços e avaliações de usuários, para auxiliar na decisão de compra.
Gemini	1	O tópico foca na culinária prática com fornos no Buscapé, mencionando receitas, dicas de uso e diferentes funcionalidades

		dos fornos, visando inspirar e facilitar o dia a dia na cozinha.
--	--	--

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
1	2664	301	534

Tópico	Entidades frequentes	Frequência
1	brastemp, samsung, brasil, bosch, electrolux, eletrolux, britania, midea, arno, novo	76, 29, 20, 18, 16, 11, 10, 9, 8, 8

Tópico	Comentário	Sentimento
1	Na ocasião da compra estava em duvida em relação a fornos mais baratos. Entretanto valeu muito a diferença de preço pois o forno é muito bom, além de ser bonito. Utilizo o meu diariamente e estou extremamente satisfeito na medida em que substituiu totalmente o meu forno a gás. O que gostei: Várias funções que permitem a você controlar bem como aquecerá o interior do forno, mas ao mesmo tempo simples de utilizar. A estética é muito bonita e ficou muito bem na minha cozinha. A porta que abre para o lado também facilita muito para um forno embutido. O que não gostei: Nenhum até o momento.	positivo
1	Esta panificadora é muito prática, faz tudo praticamente sozinha, basta colocar os ingredientes e programar de acordo com a receita, além de fazer diversos tipos de pães ainda faz geléias e panetones. Eu adoro, aliás com ela não preciso sujar vários itens para fazer uma receita, enquanto ela trabalha muito bem, eu curto a minha família. O que gostei: Prática, econômica,	positivo

	resistente e prepara diversas receitas. O que não gostei: Sem contras, até o momento!	
1	Produto excelente. Marca confiável. Super útil e prático. Muitas vezes pensamos que é grande demais para uma família pequena, mas é ideal, pois dá pra preparar qualquer tipo de prato e principalmente, não te deixa na mão em ocasiões que a casa estiver cheia. Muito bom mesmo, uso e recomendo sempre. O que gostei: Muito bom mesmo. Fácil de ser utilizado, higienizado e muito prático. O que não gostei: Ainda não há nada contra.	positivo
1	seria um excelente produto se nao fosse a rapida formacao de placas de gelo. e olha que o meu esta na temperatura minima O que gostei: este freezer atendeu as minhas expectativas, quanto ao armazenamento. o congelamento é extremamente rapido; mto bom para quem deseja congelar alimentos rapidamente. O que não gostei: infelizmente, nao estou satisfeita quanto a produçao de gelo que é enorme. as paredes do freezer produzem placas de gelo, assustadoramente, isto prejudica os alimento ali depositados	neutro
1	O PONTO POSITIVO GERAL É O ATENDIMENTO A CLIENTE DO PONTO FRIO E DA LG QUE ESTÃO EM CONTATO DIRETO TENTANDO RESOLVER AS DISPARIDADES. O que gostei: O SISTEMA DE PROGRAMA É BOM. O que não gostei: DIMENSÕES INTERNAS NÃO SÃO DE 30 LITROS, NA VERDADE CONSIDERANDO OS ABAULADOS CHEGA NO MÁXIMO A 27 LITROS NÃO É INOX,É PINTURA SOBRE PLÁSTICO NA PORTA IMITANDO INOX. POTENCIA É 900W, NÃO 1000W. VEIO COM DEFEITO NO TIMER NÃO FUNCIONA. DÁ PARA ASSAR UM FRANGO PEQUENO. ÁREA UTIL É 21L, C/ O PRATO.	neutro

1	Gostaria que todos os refrigeradores fossem fundos como este. Geralmente, os fabricantes preferem fazer o refrigerador largo, mas pouco fundo, o que dificulta o armazenamento de panelas fechadas. O que gostei: Ele é um pouco mais estreito que o habitual, mas bem mais fundo, o que facilita muito para guardar panelas fechadas. Inclusive, cabe folgada, na primeira prateleira, uma panela de pressão de 7 litros. O que não gostei: O material de que é feito as gavetas é muito frágil. Isso se aplica também à tampa da gaveta de legumes.	neutro
1	Compramos duas. Ambas apresentaram o mesmo defeito em pouco tempo (seção de baixo não refrigera). Uma quebrou antes de um ano de uso e foi consertada, pois tinha garantia estendida da loja. A outra quebrou em 27 meses. O conserto é tão caro que não vale a pena. O que gostei: As duas divisões com temperaturas diferentes. O que não gostei: Pouca durabilidade: apresentou defeitos sérios muito rápido.	negativo
1	Comprei o produto acreditando na marca, tratando o acabamento não é tão perfeito assim. Os 3 botões de girar parecem ser mal encaixados na estrutura, a luzinha vermelha apresenta "rebarbas" do plástico. O forno é ótimo, o que pega é só o acabamento que não é tão caprichado assim para uma marca famosa como a Suggar.	negativo
1	o produto foi caro e veio quebrado fiquei muito tempo passando calor graças ao ponto frio e eletrolux, falta de atenção por parte do ponto frio O que não gostei: o produto foi caro e veio quebrado fiquei muito tempo passando calor graças ao ponto frio e eletrolux	negativo

A seguir, é apresentada a aplicação completa do método para o tópico 2, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
2	[('perfume', 0.73976743), ('fragancia', 0.63734144), ('fragrancia', 0.62524986), ('aroma', 0.62431246), ('cheiro', 0.57979614), ('produto', 0.18646605), ('sensacao', 0.14648572), ('preco', 0.14017032), ('fixacao', 0.12782669), ('elegancia', 0.122545585)]	3458

Modelo de Linguagem	Tópico	Frase
ChatGPT	2	O tópico trata sobre a análise de perfumes, fragrâncias e aromas em produtos de beleza na plataforma de comércio eletrônico Buscapé.
ChatGPT	2	Este tópico aborda a avaliação de perfumes, fragrâncias e aromas em produtos de beleza na plataforma Buscapé.
ChatGPT	2	Trata-se da análise de perfumes, fragrâncias e aromas de produtos de beleza na plataforma de comércio online Buscapé.
Gemini	2	O tópico trata sobre perfumes e fragrâncias no Buscapé, com foco em aromas, sensações, fixação, elogios e dicas de escolha.
Gemini	2	O tópico aborda a experiência sensorial com perfumes no Buscapé, incluindo aromas marcantes, fixação duradoura,

		elogios recebidos e sugestões para diferentes ocasiões.
Gemini	2	O tópico foca na busca pelo perfume ideal no Buscapé, mencionando fragrâncias para diversos gostos, dicas de especialistas e avaliações de usuários, visando a satisfação e bem-estar dos clientes.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
2	3116	212	130

Tópico	Entidades frequentes	Frequência
2	brasil, dior, boss, hugo, klein, calvin, ferrari, black, jean, carolina	44, 26, 20, 19, 18, 16, 16, 11, 9, 8

Tópico	Comentário	Sentimento
2	fragrância leve e fixação duradoura.. adoroooo esse perfume	positivo
2	Adorei o perfume tem uim cheiro muito gostoso e uma fixação maravilhosa. O que gostei: Excelente aroma e fixação. O que não gostei: não tem	positivo
2	boa fixação, boa fragrancia,boa embalagem,bom preço O que gostei: boa fixação O que não gostei: nenhuma	positivo
2	Gostei da frasco, mais a fragrância foi mais ou menos, um perfume para o dia-a-dia O que gostei: bom preço, um perfume para o dia-adia O que não gostei: fixação na minha pele não dura muito, a fragrância não me agradou muito	neutro

2	É bom, mas eu não volto a comprar. O que gostei: Aroma é agradável, porém a fixação deixa a desejar. O que não gostei: Péssima fixação. Pelo preço deveria ser melhor. Não deixo de comprar Jean Paul Galtier ou Angel por outro desconhecido nunca mais.	neutro
2	Perfume muito cheiroso, doce sem ser enjoativo, mas como a fixação é ruim, creio que só valha a pena adquirir em preço de promoção. O que gostei: fragrância muito agradável, cheiro incomum, embalagem bonita, promoção muito boa. O que não gostei: Pouca fixação, a tampa do frasco começou a descascar já nos primeiros usos.	neutro
2	o perfume é ate bonzinho quando coloca, alguns segundos depois muda o cheiro, e a fixação deixa a desejar O que gostei: preço, o cheiro + ou- , mas pra usar apos banho dá. O que não gostei: fixação	negativo
2	Olá a todos. Comprei o coffret que está na promoção na Sack's. Chegou dentro do prazo e as amostras grátis vieram certinhas como escolhi. Ou seja, nota dez pra Sack's. Agora quanto ao Guerlain Homme Masculino, achei uma fragrância muito fraca. Tenho que colocar muito para senti-lo presente em mim e com isso não dura nada. Acho indicada para pessoas que gostam de fragrâncias discretas, suaves do tipo que só quando se chega perto é que se sente, pois não é de deixar rastro. É uma fragrância boa cítrica amadeirada e, diga-se de passagem, é muito mais amadeirada que cítrica e com isso é super inovadora e diferente, mas sem nada demais. Se tivesse uma presença maior, eu poderia pensar em adotá-la como uma das fragrâncias que uso. Já o sabonete líquido é maravilhoso. Um abraço a todos. O que gostei: Comprei o coffret que está na	negativo

	promoção na Sack's. Chegou dentro do prazo e as amostras grátis vieram certinhas como escolhi. Ou seja, nota dez pra Sack's É uma fragrância boa cítrica amadeirada e, diga-se de passagem, é muito mais amadeirada que cítrica e com isso é super inovadora. O que não gostei: Guerlain Homme Masculino, achei uma fragrância muito fraca. Tenho que colocar muito para senti-lo presente em mim e com isso não dura nada. Acho indicada para pessoas que gostam de fragrâncias discretas, suaves do tipo que só quando se chega perto é que se sente, pois não é de deixar rastro.	
2	Apesar de ser potente, prático pra lavar, a assistência técnica no Brasil é exageradamente precária. Enviei o aparelho no terceiro mês pra uma troca de motor, porque estava exalando cheiro de queimado, e até hoje não obtive retorno. Entrei com um processo no Procon e lá fui informado que quem responderá a este será o fornecedor, no caso Carrefour. Isto porque o produto é importado. O que gostei: Muito bonito e potente. Tem funções avançadas que não se encontra no mercado brasileiro. O que não gostei: Após 3 meses de uso o aparelho começou a exalar cheiro de queimado.	negativo

A seguir, é apresentada a aplicação completa do método para o tópico 3, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
3	<pre>[('carro', 0.97605544), ('veiculo', 0.90444136), ('driver', 0.68223554), ('motor', 0.64707744), ('seguranca', 0.53408116), ('roda', 0.5220765), ('cara', 0.49798185), ('freio', 0.4972417), ('versao', 0.47281235), ('bebe', 0.47187686)]</pre>	806

Modelo de Linguagem	Tópico	Frase
ChatGPT	3	O tópico trata sobre características e especificações de carros e veículos em uma plataforma de comércio eletrônico como o Buscapé, abordando itens como motor, segurança, rodas e versões.
ChatGPT	3	Neste tópico, são analisadas as especificações e características de carros e veículos em uma plataforma de comércio eletrônico como o Buscapé, destacando elementos como motor, segurança e rodas.
ChatGPT	3	Trata-se de informações detalhadas sobre carros e veículos em uma plataforma de comércio eletrônico como o Buscapé, explorando aspectos como motor, segurança, rodas e versões disponíveis.
Gemini	3	O tópico trata sobre carros e veículos, com foco em modelos, características, segurança e diferentes versões disponíveis.
Gemini	3	O tópico aborda a compra de carros online no Buscapé, incluindo pesquisa de modelos, comparação de preços e avaliação de aspectos como segurança e motor.
Gemini	3	O tópico foca em encontrar o carro ideal no Buscapé, mencionando variedade de modelos, filtros de pesquisa, informações detalhadas e opções para diferentes

		necessidades e públicos.
--	--	--------------------------

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
3	690	55	61

Tópico	Entidades frequentes	Frequência
3	fiat, brasil, ford, siena, honda, fiesta, palio, ferrari, corsa, chevrolet	32, 13, 11, 8, 7, 7, 7, 6, 6, 5

Tópico	Comentário	Sentimento
3	É um carro de ponta, grande, com boa tração, ótimo de dirigir e com um ótimo custo benefício. Além de tudo é um carro bonito e muito "família". O que gostei: Conforto, beleza e desempenho. O que não gostei: Preço 2. tive um classe a ano 2002 sou apaixonado por mercedes foi meu primeiro carro da mercedes logo quero comprar outro	positivo
3	tive um classe a ano 2002 sou apaixonado por mercedes foi meu primeiro carro da mercedes logo quero comprar outro	positivo
3	É um produto com design maravilhoso, confortável, bom para dirigir, pequeno e compacto, possui uma boa potência mesmo sendo 1.0,não consome muito combustível, possui manutenção fácil e não muito cara, sem contar na segurança que ele oferece ou seja é produto que qualquer pessoa gostaria de ter inclusive eu que estou pensando em ter um.Tanto que recomendo a qualquer pessoa a adquiri-lo. O que gostei: Um carro que possui um design chamativo é economico, confortável, ótimo para dirigir lhe	positivo

	proporciona um conforto excelente, além ser muito bonito. O que não gostei: único pró contra é não ser de quatro de portas, mas o que não é tão significativo.	
3	a moto é muito boa, resistente, se não fosse pela bateria seria perfeita. mas tem que falar pra criança não fazer curva fechada em velocidade se não a moto vira. Isso vai de cada pai ou mãe. O que gostei: a moto é muito boa,linda e resistente a crianças peraltas como é meu neto pneus resistentes. Adorei os adesivos, buzina, assento, rodas etc.. O que não gostei: a bateria não é legal, comprei 2 motos, e logo no começo a 1 apresentou defeito na bateria que foi trocada pela garantia. E agora sem garantia a outra parou de funcionar, vou ter que desembolsar para consertar e o pior é que aonde conserta é em outra cidade no meu estado. E tem também os retrovisores que se bater eles são frágeis e desmontam	neutro
3	Boa maquina mais depois um ano uso ele tem varias problemas com as partes plásticos. O motor e boa, as rodas de plástico não e o mesmo qualidade e o partes para protegi as rodas a frente e fraco demais. O que gostei: Prós: O tracção nas rodas dianteiras é ótimo, super fácil para cortar grama. Não canse muito, eu uso o maquina para o chácara de 5000 m2 O que não gostei: Contras: - qualidade das rolamentos das rodas (melhor e rolamento de esferas) e o proteção plastico das rodas dianteiras e fraco de mais. Parcial um boa maquina mais varias details e a baixa o nível certo. 3. Por ser o pacote JOY não tem muitos itens que poderia vir já de série como DH e VE. O que gostei: motor O que não gostei: itens de série	neutro
3	Por ser o pacote JOY não tem muitos itens que poderia	neutro

	vir já de série como DH e VE. O que gostei: motor O que não gostei: itens de série	
3	so tem preço e um produto pessimo,carro q custa 6 mil dolaras la fora custa 25.000,00 a q O que gostei: era p ser um bom carro O que não gostei: n tem conforto nem um	negativo
3	UMA VERDADEIRA FALTA DE RESPEITO COM O CONSUMIDOR PRODUTO NÃO PODERIA SER COMERCIALIZADO. O que gostei: VISUAL O que não gostei: PESADA DEMAIS ,SELM DESTROI QUALQUER RABO,PNEUS E RODAS DE TERCEIRA ENFIM UMA PESSIMA AQUISIÇÃO...	negativo
3	Apesar da economia de ser 1.0, do conforto e design do carro, ele peca por ter um motor rocan importando. Uma vez com problema voce pagara caro pelos ajustes!	negativo

10.4. B2W DIGITAL

O método foi aplicado ao conjunto de dados da plataforma de comércio eletrônico B2W Digital, que foi disponibilizado pela empresa Americanas em seu repositório do GitHub. Esse conjunto abrange aproximadamente 132 mil avaliações coletadas entre Janeiro a Maio de 2018. Durante a análise desse conjunto de dados, foi observado que várias avaliações não continham comentários, consistindo apenas na atribuição de um número de estrelas (variando de 1 a 5). Portanto, foram consideradas apenas as avaliações que continham comentários válidos, resultando em um total de 129.863 comentários.

Para que fosse possível ter uma demonstração visual das palavras mais mencionadas nos comentários das avaliações foi gerada uma nuvem com tais palavras, onde é possível visualizar que palavras como “produto”, “prazo”, “qualidade” e “dia” foram frequentemente escritas.

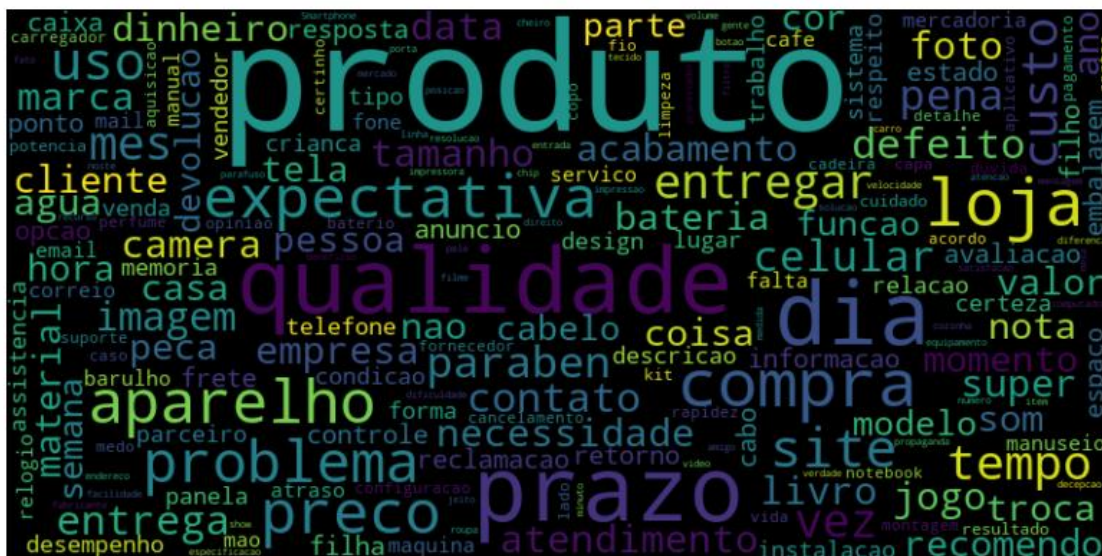


Figura 24 - Nuvem de palavras dos comentários das avaliações da B2W Digital. Fonte: A autora.

Prosseguindo com a modelagem de tópicos, foram identificados 63 tópicos, sendo que somente os 4 primeiros foram escolhidos para análise.



Figura 25 - Gráfico de barras dos 8 tópicos mais frequentes do conjunto de dados da B2W Digital. Fonte: A autora.

A seguir, é apresentada a aplicação completa do método para o tópico 0, incluindo as frases descritivas associadas. Também são mostrados os resultados da análise de sentimentos, com a quantidade de comentários classificados como “positivos”, “negativos” e “neutros”. Além disso, são listadas as entidades nomeadas encontradas e suas frequências. Por fim, são exibidos os três primeiros comentários

relacionados ao tópico, categorizados por sentimento.

Tópico	Palavras do Tópico	Frequência
0	[('expectativa', 0.6213931), ('prazo', 0.5424142), ('produto', 0.46128765), ('tempo', 0.43545085), ('qualidade', 0.40994468), ('atraso', 0.39468712), ('momento', 0.32398754), ('preco', 0.32250255), ('desempenho', 0.29777175), ('hora', 0.29508886)]	61336

Modelo de Linguagem	Tópico	Frase
ChatGPT	0	O tópico trata sobre a expectativa de prazo e qualidade dos produtos na plataforma de comércio eletrônico B2W.
ChatGPT	0	Este tópico aborda a análise da expectativa de prazo e qualidade dos produtos na plataforma B2W.
ChatGPT	0	O tema central é a consideração da expectativa de prazo e qualidade dos produtos oferecidos na plataforma de comércio eletrônico B2W.
Gemini	0	O tópico trata sobre a experiência do cliente com prazos de entrega no B2W, com foco em expectativas, atrasos, qualidade dos produtos e comunicação ao cliente.
Gemini	0	O tópico aborda a gestão de prazos de entrega no B2W, incluindo prazos realistas, comunicação clara com o cliente e otimização do processo logístico, visando a satisfação do cliente.

Gemini	0	O tópico foca na entrega eficiente no B2W, mencionando prazos curtos, atualizações frequentes do status do pedido e cumprimento das expectativas do cliente, para garantir uma ótima experiência de compra.
--------	---	---

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
0	42995	2698	15643

Tópico	Entidades frequentes	Frequência
0	americanas, americana, samsung, brasil, motorola, sao, philco, correios, sony, eletro	947, 795, 228, 124, 95, 64, 51, 51, 45, 44

Tópico	Comentário	Sentimento
0	Bem macio e felpudo...recomendo. Preço imbatível e entrega rápida. Compraria outro quando precisar	positivo
0	Produto excepcional! recomendo!!! inovador e gostoso de ler!!	positivo
0	Além de higiênico tem o tamanho ideal. Só faltou colher e garfo de sobremesa. Talvez se viesse com mais peças, eu gostaria mais.	positivo
0	Ué não entendi! Estava procurando recomendações a respeito da pipoqueira da Cadence vi esses comentários . Agora fui ver a da Britânia: os mesmos comentários!!!!?? Como assim???	neutro
0	Quero saber se essa smart TV tem o Netflix ou se dá pra baixar	neutro

0	se vc tem um carro antigo, sem bluetooth, mas com entrada auxiliar, vc precisa comprar um destes!	neutro
0	Não compra este produto, pois o mesmo veio com os botões travados e ao tentar ligar o mesmo nem ligou. Péssima qualidade.	negativo
0	Uma enganação! Esperei quase um mês e cancelaram o pedido. E o produto continua disponível. Nem entraram em contato. Palhaçada!!! #carrodemola	negativo
0	Na estação do som tudo correto mas a estalar os autofalante não sai som.	negativo

A seguir, é apresentada a aplicação completa do método para o tópico 1, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
1	[('celular', 0.77535903), ('camera', 0.7596413), ('smartphone', 0.6299713), ('telefone', 0.6159553), ('android', 0.49429342), ('foto', 0.46947214), ('imagem', 0.46369377), ('aparelho', 0.45163262), ('video', 0.4072774), ('fone', 0.40014157)]	8320

Modelo de Linguagem	Tópico	Frase
ChatGPT	1	O tópico trata sobre análises de dispositivos móveis, especificamente celulares e smartphones com foco em câmeras e funcionalidades de imagem na plataforma de comércio eletrônico B2W.
ChatGPT	1	Este tópico aborda análises de dispositivos móveis, especialmente celulares e smartphones, com ênfase em câmeras e

		funcionalidades de imagem, na plataforma B2W.
ChatGPT	1	O tema central é a avaliação de celulares e smartphones, com destaque para câmeras e recursos de imagem, na plataforma de comércio eletrônico B2W.
Gemini	1	O tópico trata sobre celulares e smartphones no B2W, com foco em modelos, câmeras, qualidade das fotos e vídeos, e sistema operacional Android.
Gemini	1	O tópico aborda a escolha do celular ideal no B2W, incluindo marcas, funcionalidades, preços e avaliações de usuários, para auxiliar na decisão de compra.
Gemini	1	O tópico foca na experiência mobile no B2W, mencionando variedade de celulares, recursos avançados, conectividade e praticidade no dia a dia, visando atender às necessidades dos clientes.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
1	5545	503	2272

Tópico	Entidades frequentes	Frequência
1	samsung, americanas, americana, android, motorola, asus, moto, galaxy, brasil, zenfone	303, 166, 130, 128, 107, 88, 60, 46, 43, 37

Tópico	Comentário	Sentimento
1	Bom custo benefício. Adequado para pessoas que fazem um uso casual do aparelho . A máquina fotográfica é excelente . A tela fornece imagens ótimas . Os recursos do aparelho são fáceis de manejar.	positivo
1	Adorei o smartphone, excelente acabamento, veloz e muito intuitivo	positivo
1	Aparelho muito bom! Ótima imagem e fotos maravilhosas.	positivo
1	Eu comprei este produto, entretanto me foi entregue um produto diferente, foi entregue um produto que aceita somente um chip (modelo k33b37) além disso veio faltando o fone. Estou em comunicação com a parceira Eletrobest e eles já me enviaram o código postal para eu enviar e fazer a troca. O produto que veio, apesar de errado , é um produto bom, design bonito, leve, com muitas funcionalidades, entretanto como já mencionado, a empresa parceira cometeu dois erros que eu considero erros muito primários e graves (mandar o produto errado e faltando o fone)	neutro
1	Um aparelho muito bom mas com alguns detalhes muito ruim também. Minha critica é que o aparelho não tem a opção de ocultar meu numero . O fone de ouvido é ridiculo por um valor alto do aparelho. Mas resumindo é um bom aparelho.	neutro
1	Pelo valor do aparelho, ele entrega o que promete. É bonito, câmera razoável. Recomendo	neutro
1	Balcão bonito. Entrega no prazo. Material frágil. Veio faltando parafusos para fixação do tampo, assim como faltou a proteção (cor preta) atrás do tampo, que aparece na imagem. Montagem fácil, mas o manual de	negativo

	instruções poderia ser mais completo. Não recomendo pela falta dos parafusos e da proteção.	
1	Comprei um faz um ano e dois meses, já foi a assistência 2 vezes com o mesmo defeito e o problema não foi solucionou. Smartphone horrível!	negativo
1	Quando o telefone chegou, aqui na minha cidade ja estava quase 200.00 mais barato	negativo

A seguir, é apresentada a aplicação completa do método para o tópico 2, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
2	[('livro', 0.778964), ('leitura', 0.49999076), ('leitor', 0.47706303), ('explicacao', 0.3518085), ('filme', 0.3455627), ('papel', 0.32739508), ('estudo', 0.32577693), ('folha', 0.31732905), ('desenho', 0.3047537), ('historia', 0.29821855)]	2072

Modelo de Linguagem	Tópico	Frase
ChatGPT	2	O tópico trata sobre livros e leitura, com ênfase em experiências de leitura, leitores e exploração literária na plataforma de comércio eletrônico B2W.
ChatGPT	2	Este tópico aborda livros e leitura, com destaque para experiências de leitura, leitores e exploração literária, na plataforma B2W.
ChatGPT	2	O tema central é a análise de livros e leitura, com foco em experiências de leitura, leitores e exploração literária, na

		plataforma de comércio eletrônico B2W.
Gemini	2	O tópico trata sobre livros e leitura no B2W, com foco em variedade de títulos, gêneros literários, autores e promoções para incentivar a leitura.
Gemini	2	O tópico aborda a experiência do leitor no B2W, incluindo facilidade na busca por livros, avaliações de outros leitores e diferentes formatos de leitura, como e-books e audiobooks.
Gemini	2	O tópico foca no universo dos livros no B2W, mencionando lançamentos, novidades literárias, dicas de leitura e promoções imperdíveis para os amantes de livros.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
2	1723	120	229

Tópico	Entidades frequentes	Frequência
2	americana, americanas, brasil, jesus, sao, julia, quinn, mala, augusto, paulo	23, 21, 20, 15, 9, 7, 7, 7, 7, 6

Tópico	Comentário	Sentimento
2	Livro excelente! A entrega feita muito rápida. Super recomendo!	positivo
2	O livro é muito bom, recomendo, chegou no tempo certo.	positivo

2	Livro muito bom, entregaram antes do prazo e em ótimo estado.	positivo
2	Ainda estou estudando o livro e prefiro não fazer comentarios, Porém em uma primeir e superficial análise acho o livro bem estruturado .	neutro
2	Esta edição limitada da Darkside ficou linda, o livro é incrível, porém, o que não me agradou foi a falta de cuidado pelo qual teve a vendedora T2 Shop para entrega do meu produto. Nunca tinha comprado na mesma e possivelmente nem voltarei a comprar e indicar para meus amigos leitores. Meu livro chegou cum um amassado bem feio no canto inferior direito, tanto na parte da frente quanto na parte de trás. Pelo fato de ser capa dura, consegui amenizar um pouco a situação mas ainda é bem visível, parece que ele foi jogado ou caiu ao chão de mal jeito e acabou danificado. Como era um item de colecionador, eu gostaria de ter um cuidado especial com o produto e essa situação me deixou bem chateado.	neutro
2	As histórias são legais, mas são fraquinhas...E o melhor é o primeiro livro.	neutro
2	Meu pedido foi enviado para outro local e chegou na minha casa outros livros. Realizei a devolução e o livro que eu comprei chegou numa data passada da que me interessaria para o exame do Colégio do meu sobrinho. Não abri a embalagem para não ter problemas na devolução do livro e do dinheiro ou vale.	negativo
2	Infelizmente não poderei avaliar este livro, pois ainda não o recebi, apesar de o prazo estabelecido para a sua entrega ter vencido desde o dia 05/05.	negativo
2	Comprei esse produto .ta atrasado .ate agora nao	negativo

	chegou .e nao consigo falar e fazer reclamação nessa pagina .e agora como eu fiko ?????	
--	---	--

A seguir, é apresentada a aplicação completa do método para o tópico 3, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
3	[('filha', 0.7156216), ('produto', 0.7045338), ('filho', 0.5760075), ('pai', 0.53553396), ('sobrinho', 0.53058606), ('crianca', 0.48742378), ('familia', 0.4856708), ('mae', 0.42722207), ('bebe', 0.40316063), ('entrega', 0.39172074)]	1781

Modelo de Linguagem	Tópico	Frase
ChatGPT	3	O tópico trata sobre produtos voltados para crianças e família na plataforma de comércio eletrônico B2W.
ChatGPT	3	Este tópico aborda produtos destinados a crianças e famílias na plataforma B2W.
ChatGPT	3	O tema central é a análise de produtos voltados para crianças e família na plataforma de comércio online B2W.
Gemini	3	O tópico trata sobre presentes para crianças e bebês no B2W, com foco em produtos para filhas, filhos, sobrinhos, afilhados e necessidades de pais e mães.
Gemini	3	O tópico aborda a compra de presentes infantis no B2W, incluindo variedade de produtos, faixas etárias, marcas confiáveis e promoções para todas as ocasiões.

Gemini	3	O tópico foca no universo infantil no B2W, mencionando brinquedos, roupas, itens de higiene, livros e tudo o que os pequenos precisam para se divertir e crescer com segurança, proporcionando alegria para toda a família.
--------	---	---

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
3	1460	46	275

Tópico	Entidades frequentes	Frequência
3	americanas, americana, theinsidersbrasil, samsung, lego, sao, hasbro, barbie, odin, mari	19, 13, 9, 3, 3, 2, 2, 2, 2, 2

Tópico	Comentário	Sentimento
3	Como farei aniversário dia 3 de fevereiro pedi estas panelas pro meu marido. AMEI.	positivo
3	Meu filho amou, a hélice não veio a amarela veio branca mas, tirando isso tudo ok . Gostei	positivo
3	Muito fofo e fácil de usar. Meu filho adorou. É ótimo para treinamento.	positivo
3	e ai filhod d putaa sdd sasdddddddddddaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	neutro
3	achei a história muito extensa. No meu caso, que minha filha tem 2 anos, ela não tem paciência para contar até no final. E ela fala que o lagarto é um jacaré! rs	neutro

3	Gostamos do produto , mas tive uma péssima experiência com na entrega , fiz uma compra pro chegar antes do Natal prq era um presente para minha filha , mas chegou bem depois , estou decepcionado..	neutro
3	Produto de qualidade, mas pequeno. Minha filha não tem nem um ano e meio e já quase não cabe nele! O colchonete é bem fino tbm.	negativo
3	Meu filho ainda nao tem coordenacao apurada e com isso, já danificou o aparelho. Achei bem fraco, pra ser indicado acima de 3 anos . Nao vai durar uma semana em casa	negativo
3	pedi o hoverboard dia 05/11/2017 e até hoje não chegou,liguei várias vezes,fui várias vezes nas lojas americanas falar sobre este produto e nada, é só golpe,não comprem pois é possível não chegar no prazo como o meu,e pagar um alto preço no frete e não chegar,minha filha ficou sem presente de natal e muito decepcionada,espero que chegue pelo menos NESTE ANO já que em 2017 não chegou espero tê-lo em 2019 haha	negativo

10.5. HOTÉIS BRASILEIROS

O método foi aplicado ao conjunto de dados de hotéis brasileiros obtidos da plataforma TripAdvisor, que foi disponibilizado por uma pesquisa da Universidade Federal de Viçosa. Esse conjunto abrange aproximadamente 730 mil avaliações coletadas entre fevereiro e março de 2018. Durante a análise desse conjunto de dados, foi observado que várias avaliações não continham comentários, consistindo apenas na atribuição de um número de estrelas (variando de 1 a 5). Portanto, foram consideradas apenas as avaliações que continham comentários válidos, resultando em um total de 686.741 comentários.

Para que fosse possível ter uma demonstração visual das palavras mais mencionadas nos comentários das avaliações foi gerada uma nuvem com tais

palavras, onde é possível visualizar que palavras como “hotel”, “café”, “localização” e “atendimento” foram frequentemente escritas.



Figura 26 - Nuvem de palavras dos comentários das avaliações dos hotéis brasileiros. Fonte: A autora.

Prosseguindo com a modelagem de tópicos, foram identificados 134 tópicos, sendo que somente os 4 primeiros foram escolhidos para análise.

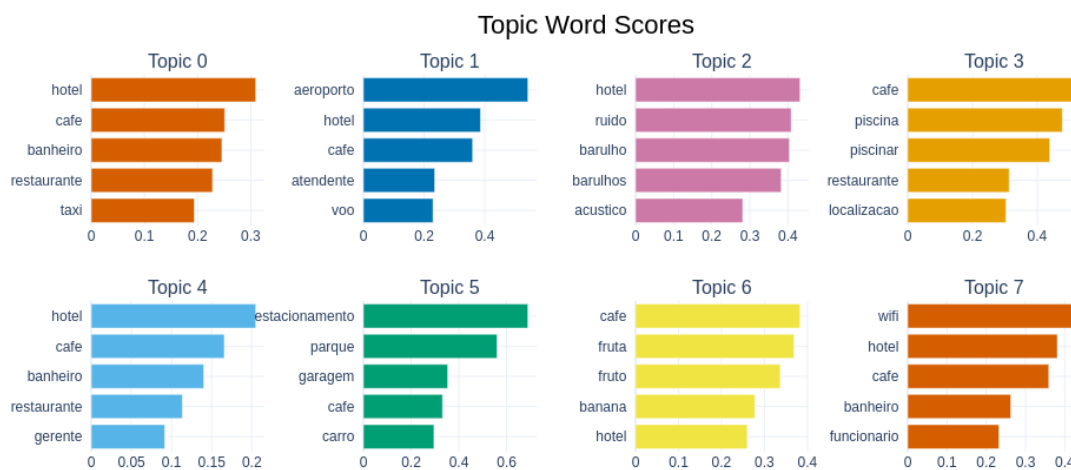


Figura 27 - Gráfico de barras dos 8 tópicos mais frequentes do conjunto de dados dos hotéis brasileiros. Fonte: A autora.

A seguir, é apresentada a aplicação completa do método para o tópico 0, incluindo as frases descritivas associadas. Também são mostrados os resultados da análise de sentimentos, com a quantidade de comentários classificados como

“positivos”, “negativos” e “neutros”. Além disso, são listadas as entidades nomeadas encontradas e suas frequências. Por fim, são exibidos os três primeiros comentários relacionados ao tópico, categorizados por sentimento.

Tópico	Palavras do Tópico	Frequência
0	[('hotel', 0.30927247), ('cafe', 0.2509985), ('banheiro', 0.24568461), ('restaurante', 0.22806373), ('taxi', 0.19381991), ('hospedagem', 0.13921237), ('quarto', 0.13642177), ('funcionario', 0.13217625), ('atendente', 0.13111483), ('custo', 0.11922985)]	215903

Modelo de Linguagem	Tópico	Frase
ChatGPT	0	O tópico trata sobre as características e serviços oferecidos por hotéis brasileiros, incluindo café da manhã, banheiro, restaurante, acesso a táxi, hospedagem, quartos, funcionários e custo.
ChatGPT	0	Este tópico aborda as amenidades e serviços disponíveis em hotéis brasileiros, como café da manhã, banheiro, restaurante e custo.
ChatGPT	0	O tema central é a análise dos serviços e facilidades oferecidos por hotéis no Brasil, como café da manhã, banheiro, restaurante e custo.
Gemini	0	O tópico trata sobre experiências em hotéis brasileiros, com foco em acomodações, café da manhã, banheiro, restaurante, atendimento e custo-benefício.
Gemini	0	O tópico aborda serviços e comodidades

		oferecidos por hotéis no Brasil, incluindo tipos de quartos, café da manhã, restaurantes, serviços de transporte e equipe de atendimento.
Gemini	0	O tópico foca na avaliação de hotéis brasileiros, mencionando acomodações, café da manhã, instalações, serviços e custo-benefício, para auxiliar na escolha do hotel ideal para sua viagem.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
0	180247	17381	18275

Tópico	Entidades frequentes	Frequência
0	sao, rio, paulo, ibis, paulista, curitiba, brasilia, negra, janeiro, brasil	6873, 5609, 5135, 4335, 4115, 3575, 3352, 3095, 2828, 2429

Tópico	Comentário	Sentimento
0	O hotel é muito bom. O atendimento é excelente. O quarto é aconchegante, os móveis não são novos mas estão conservados. Possui 2 elevadores mas na data em que fiquei no hotel só um estava funcionando. Hotel muito bom.	positivo
0	Os hotéis da Rede Bittar em Brasília servem ao seu propósito: Economia e baixo custo. Não esperem quartos luxuosos, espaçosos e mobília nova. Café da manhã decente, padrão em toda rede hoteleira, mesmo assim, melhor que os da rede Ibis.	positivo
0	Reserva sem complicações no site do hotel.	positivo

	Apartamentos reformados. Café da manhã com boa variedade de itens. O restaurante atende à noite com boas opções no cardápio, inclusive aceitando adaptações na composição do prato (pedi para substituir o arroz por salada), foram muito gentis e atenciosos. Localização excelente, vizinho ao Brasília Shopping. Setor de reservas muito atencioso.	
0	O hotel é bem localizado e silencioso, o que conta muito. Porém, deixou a desejar quanto ao café da manhã, que achei pouco variado, sobretudo para quem faz dieta com baixo consumo de farinha de trigo. Nas acomodações, o banheiro não tinha, dentro do box, nichos para se colocar o sabonete. A conexão wifi caía sempre.	neutro
0	O hotel não tem fachada bonita. Tem excelente localização. Próximo a dois shoppings, torre, estádio mané guarrincha. Quarto com armário para roupa bom , boas gavetas, escrivaninha, criados bons e bom espaço. WC bom. Enfim tamanho do quarto muito bom. Café da manhã bom. Limpeza todos os dias porém não trocam toalhas. Ar antigo, muito barulhento. sacada faltando limpeza. Vaso sanitário encardido.	neutro
0	O hotel é próximo ao Brasília Shopping, mas o hotel é um pouco velho, o Wi-Fi paga 10 reais ao dia, entra cracudo na recepção a noite e os funcionários não são simpáticos. Bom pra passar poucos dias. Não posso reclamar da limpeza, pois foi limpo todos os dias.	neutro
0	Hotel antigo, mal conservado e mal reformado, apesar de uma estadia de apenas um dia, não gostei desse hotel. O atendimento foi razoável a estrutura deixa muito a desejar, pelo preço que cobra existem opções de custo beneficio muito melhores em Brasília! Não	negativo

	voltaria me hospedar nesse hotel.	
0	Restaurante velho, sujo, encontrei barata no quarto, atendimento ruim, café da manhã com produtos vencidos. Estacionamento é público Impossível se hospedar neste hotel. Péssima qualidade de sono e experiência.	negativo
0	Fiz a reserva e quando cheguei encontrei o hotel fechado para reformas, nenhum tipo de aviso foi dado quando da realização da reserva. Tive que procurar outro hotel da rede para só então ser informado que as reservas de Bittar Plaza Hotel seriam direcionadas para outro hotel da rede. Não recomendo!	negativo

A seguir, é apresentada a aplicação completa do método para o tópico 1, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
1	[('aeroporto', 0.5422682), ('hotel', 0.38601196), ('cafe', 0.3600198), ('atendente', 0.23507735), ('voo', 0.229696), ('taxi', 0.21010327), ('hospedagem', 0.20476946), ('porto', 0.19716029), ('restaurante', 0.1946769), ('estadia', 0.18462765)]	19442

Modelo de Linguagem	Tópico	Frase
ChatGPT	1	O tópico trata sobre a conveniência de hotéis próximos a aeroportos, com destaque para serviços como café da manhã, atendimento, transporte e estadias.
ChatGPT	1	Este tópico aborda a localização

		estratégica de hotéis próximos a aeroportos no Brasil, destacando serviços como café da manhã, atendimento e transporte.
ChatGPT	1	O tema central é a conveniência oferecida por hotéis localizados próximos a aeroportos no Brasil, com foco em serviços como café da manhã, atendimento e transporte.
Gemini	1	O tópico trata sobre logística de viagens com hotéis próximos a aeroportos, focando em acomodação, café da manhã, traslados, voos, táxis e otimização do tempo do viajante.
Gemini	1	O tópico aborda facilidade de acesso a aeroportos a partir de hotéis, incluindo opções de transporte, tempo de deslocamento, serviços de recepção e café da manhã para atender às necessidades dos viajantes.
Gemini	1	O tópico foca na experiência completa de viagem com hotéis em localizações estratégicas, mencionando proximidade de aeroportos, praticidade para voos, variedade de serviços e comodidades para garantir uma estadia tranquila e eficiente.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
1	16716	1435	1291

Tópico	Entidades frequentes	Frequência
1	congonhas, santos, dumont, porto, rio, ibis, alegre, sao, paulo, bis	1778, 1741, 1477, 994, 941, 908, 890, 810, 663, 500

Tópico	Comentário	Sentimento
1	bom Hotel , um Ótimo Café da manhã só a internet é cobrada o que achamos um absurdo mais no mais fora está cobrança eu indico próximo de tudo " shopping , ônibus direto do aeroporto passa próximo , próximo ao Senado a camara federal aos prédios importantes na capital do Brasil	positivo
1	hotel com quartos renovados, amplos e bem equipados. Agilidade no atendimento. Boa localização. Ótimo custo benefício. Hotel apenas para dormir porque não há maior estrutura. Restaurante peca pela falta de infraestrutura. Café bem básico. Executivo do Aeroporto passa a 20 metros.	positivo
1	Bom hotel. Excelente localização. Simples, mas limpo e com bom café da manhã. Quarto espaçoso. O ônibus expresso que faz a linha aeroporto tem ponto bem em frente da recepção do hotel. Próximo ao shopping. Wi fi não é gratuito.	positivo
1	Um razoável café da manhã. Quartos bem limpos e arejados. Bem próximo ao aeroporto. Um bom atendimento. Pernoitamos e ao amanhecer eu, minha filha e esposa seguimos para Rio Quente. Uma viagem fantástica. Vale apenas. Abraços.	neutro
1	A localização, a apenas 20 min. do aeroporto é muito bom. bom tamanho do. pequeno-almoço superior do quarto é bom. o restaurante para média de jantar. o serviço no bar é correcto, mas o check-in foi de longe	neutro

	muito lento considerando que houve poucas pessoas o check-in.	
1	O hotel parece mais luxuoso do que realmente é. Os quartos são bons, porém naa excepcional. O café da manhã é normal, porém bom. O que há de melhor é a localização, pois fica perto das Embaixadas. E fica no plano piloto.	neutro
1	Não recomendo. Hotel velho, cheiro de mofo com uma estrutura péssima. tivemos problema no o funcionário. wi-fi tinha mais não funcionava muito bem. a vantagem é que é perto do shopping e não tao longe do aeroporto.	negativo
1	Fiquei na parte das suítes, quarto grande e espaçoso com uma bela vista pra cidade, shopping a menos de 5 minutos a pé, ponto de onibus em frente ao hotel e 1 executivo que vai para o aeroporto e não tem muita opção para comer a noite próximo ao hotel.	negativo
1	Uma pena nao me permitir checkout °s 2 pm no lugar de 12 pm... sendo que vou ficar por 3 horas sem fazer nada esperando de ir ao aeroporto. A recepçao falou que tem muitos clientes chegando. Outros hoteis em mesma situação conseguiram me ajudar.	negativo

A seguir, é apresentada a aplicação completa do método para o tópico 2, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
2	[('hotel', 0.43177387), ('ruído', 0.40849078), ('barulho', 0.404002), ('barulhos', 0.38230628), ('acustico', 0.28142422), ('quarto', 0.25774026), ('som', 0.25541312), ('sala', 0.25330818),	9130

	('acustica', 0.22966036), ('janela', 0.22437355)]	
--	---	--

Modelo de Linguagem	Tópico	Frase
ChatGPT	2	O tópico trata sobre a qualidade do ambiente acústico em quartos de hotéis brasileiros, abordando questões relacionadas a ruído, barulho, acústica e conforto sonoro.
ChatGPT	2	Este tópico discute a experiência relacionada ao ambiente acústico em quartos de hotéis no Brasil, considerando ruído, barulho e conforto sonoro.
ChatGPT	2	O tema central é a análise da qualidade do ambiente acústico em quartos de hotéis brasileiros, abordando questões como ruído, barulho e acústica.
Gemini	2	O tópico trata sobre incômodos sonoros em hotéis brasileiros, com foco em ruídos, barulhos, acústica dos quartos, sons externos e impacto na experiência dos hóspedes.
Gemini	2	O tópico aborda dificuldades para dormir em hotéis por causa de barulhos, incluindo ruídos internos e externos, qualidade do sono dos hóspedes e medidas para garantir um ambiente tranquilo.
Gemini	2	O tópico foca na importância da acústica no conforto dos hotéis, mencionando isolamento acústico dos quartos, qualidade das janelas, ruídos do ambiente externo e ações para garantir uma boa noite de sono

		aos hóspedes.
--	--	---------------

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
2	5462	1557	2111

Tópico	Entidades frequentes	Frequência
2	sao, paulo, paulista, curitiba, ibis, rio, porto, brasilia, alegre, bis	290, 230, 225, 184, 159, 122, 116, 104, 101, 89

Tópico	Comentário	Sentimento
2	Hotel simples mas muito bom, bem localizado, perto de shopping e com fácil acesso para diversas conveniências. O problema era o ar condicionado do apartamento vizinho ao meu. Acordei na primeira noite achando que era o aparelho do meu apartamento, desliguei e o barulho continuava. Se resolverem o problema recomendo	positivo
2	Hotel bom e barato. Para quem vai para precisa visitar órgão públicos é um boa opção. Para os concurseiros o hotel fica próximo dos grandes locais de provas. A proximidade com os shoppings do eixo fundamental é outra vantagem. Quartos limpos e bom atendimento. Café da manhã agradável. A única reclamação era um barulho parecido com aparelho de ar-condicionado que era ouvido de instante em instante.	positivo
2	Fiquei no hotel no período da copa das confederações. A localização é excelente. Pode-se ir ao estádio tranquilamente a pé, próximo inclusive ao mirante da Antena de TV (fica entre o hotel e o estádio). As acomodações são simples, porém limpas. A ducha é	positivo

	agradável e a cama confortável. Café da manhã com opções. Há shopping e Mcdonalds próximo caso se procure algo rápido para alimentação. Pessoal da recepção é esclarecedor e prestativo. De ressalva apenas o ar condicionado que faz grande barulho!	
2	Bom hotel, com uma das diárias mais baratas de Brasília. Ainda assim, custou-me R\$ 455,00, duas diárias. É muito caro pelo que oferece. Tem ar-condicionado nos quartos e um bom café da manhã, mas não dispõe de internet gratuita (o serviço é terceirizado e caro, como tudo em Brasília). Além disso, pedi uma acomodação silenciosa, mas onde me colocaram, dava para ouvir o barulho de caminhões recolhendo lixo dos hotéis do Setor Hoteleiro Norte, no meio da madrugada. Há uma funcionária muito gentil, a Dona Diva.	neutro
2	O quarto é amplo e confortável, porém a higiene de alguns itens deixou a desejar. O que incomodou mesmo foi o barulho muito alto que vinha do edifício ao lado, da rua e da sirene da garagem. O isolamento do quarto não é bom. Tentei trocar de quarto, mas não havia mais quartos disponíveis em andares mais altos. No mais o hotel atende bem às necessidades de quem vem à Brasília e deseja estar bem localizado.	neutro
2	Fiquei no bloco D, vista para a torre de TV. Razoável, mas o isolamento acústico deixa a desejar, se o vizinho de quarto for barulhento, fica complicado. Atendimento bom, café da manhã bom. Cama muito boa. A internet é instável, para ter um serviço decente, tem que pagar (R\$ 30,00 por 6 horas de uso).	neutro
2	Hotel apesar de bem localizado tem uma péssima estrutura, fiquei no quarto 127, onde era notório que era	negativo

	um quarto improvisado, não tinha janela, as cortinas não impedem o sol, como este quarto é localizado para os fundos do hotel, se abrisse a cortina o quarto ficava exposto, pois nem tem insulfilme, o banheiro não tem nenhuma forma de ventilação, todo cheiro ia para o quarto, a noite o barulho do hall do hotel era de total incomodo, ao ligar a noite para recepção para saber sobre o horário de silencio me foi informado que iniciava as 00, ou seja tínhamos que esperar até a meia noite para conseguir dormir. Toda hora que tocava o telefone da recepção se ouvia desse quarto. Então, se pretende ter uma noite de sono não se hospede neste hotel. Ainda sobre o hotel, foram colocados 3 colchoes no chão, uma das "colegas de quarto" tinha dificuldade em levantar e se deitar pela baixissima altura, não tinha um minimo espaço entre os colchões. Foi a minha pior estadia em hotéis em Brasília.	
2	Fiquei hospedado no hotel por apenas um dia e não recomendo para ninguém o atendimento na recepção era péssimo. Os quartos antigos ar condicionado velho e barulhento e o frigobar na mesma condição. Tinha poeira no quarto. Café da manhã sem ter quase nenhuma opção.	negativo
2	hotel de pessima qualidade sonora, muito barulho, sem descanso nenhum, localizado em um lugar de ma qualidade, internet alem de ser paga é horrivel, cai toda momento, cafe de ma qualidade, sem diversos pratos e frios,	negativo

A seguir, é apresentada a aplicação completa do método para o tópico 3, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
3	[('cafe', 0.50803655), ('piscina', 0.4774949), ('piscinar', 0.43853724), ('restaurante', 0.31329593), ('localizacao', 0.30336785), ('local', 0.28517133), ('buffet', 0.25952744), ('lugar', 0.23144308), ('almoco', 0.22127476), ('cozinha', 0.21221772)]	5510

Modelo de Linguagem	Tópico	Frase
ChatGPT	3	O tópico trata sobre as comodidades e atrativos oferecidos por hotéis brasileiros, incluindo café da manhã, piscina, restaurante, localização privilegiada e serviços de alimentação.
ChatGPT	3	Este tópico aborda os atrativos e serviços oferecidos por hotéis no Brasil, como café da manhã, piscina, restaurante e localização privilegiada.
ChatGPT	3	O tema central é a análise das comodidades disponíveis em hotéis brasileiros, destacando café da manhã, piscina, restaurante e localização estratégica.
Gemini	3	O tópico trata sobre estrutura de lazer e alimentação em hotéis brasileiros, com foco em café da manhã, piscina, restaurante, buffet, localização e opções de almoço.
Gemini	3	O tópico aborda comodidades oferecidas por hotéis para o lazer e bem-estar dos hóspedes, incluindo café da manhã,

		piscina, restaurante, buffet e localização estratégica.
Gemini	3	O tópico foca na experiência gastronômica e de lazer em hotéis, mencionando café da manhã, piscina, restaurante, buffet, variedade de opções e qualidade dos serviços para atender às necessidades dos hóspedes.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
3	4907	282	321

Tópico	Entidades frequentes	Frequência
3	brasilia, sao, rio, careca, paulo, negra, brasil, paulista, natal, salvador	67, 66, 52, 51, 50, 44, 42, 42, 39, 38

Tópico	Comentário	Sentimento
3	Excelente localização com vista para a esplanada dos ministérios, atendimento muito bom, quarto confortável e funcional. O café da manhã é completo mas senti falta de uma variedade maior de pães. Academia bem completa. Piscina na cobertura com 1 M de profundidade.	positivo
3	Excelente conjunto de flat na região central de Brasília. Fica ao lado do Conjunto Nacional. Oferece piscina , academia, hidromassagem, restaurante e embaixo há algumas opções de comércio. Muito limpo. Vale a pena.	positivo
3	Ótima localização, localizado no setor hoteleiro norte conta com um bom serviço. O café da manhã é bem	positivo

	completo. Possui uma área externa comum com uma pracinha bem aconchegante com bares e restaurantes, possui uma vista incrível do terraço e uma boa piscina além da academia que é bem completa	
3	Dois pontos precisam de uma reavaliação: considerei a comida servida no jantar ruim, e a relação custo x benefício = cara. No quarto em que me hospedei 1218, o ralo do box de banho estava entupido, o que gerou desconforto, visto que a água enchia o local, parecendo uma mini piscina . É necessário observar a manutenção periódica desses detalhes. No mais o atendimento dos funcionários foi excelente, o quarto era muito charmoso e bem decorado. A vista maravilhosa, o nível de ruído mínimo. Retornarei quando precisar, com toda a certeza. ObrigadaMaria Isabel Alves Dumaresq	neutro
3	O quarto é muito bom, mas as instalações de lazer deixaram a desejar. A piscina estava fechada e a academia sinceramente precisa melhorar muito. Minha sugestão é inclusive deixá-la fechada enquanto não colocam novos aparelhos, fica menos feio.	neutro
3	Quarto arrumado e limpo, td bem organizado, com poucas opções no frigobar. A parte mais frustrante foi a piscina , muito suja, inutilizável. Fiquei hospedado em um domingo e segunda de manhã não consegui usar a piscina . Café da manhã bom, mas sem grandes atrativos. Custo benefício, bom.	neutro
3	O prédio e os apartamentos são agradáveis, exceto pela "piscina " do terraço, que deveria ser renomeada "banheira", ou pela TV de 1990. Mas serviço é ruim, muito ruim. A instalação de Wi-fi é uma piada. Qualquer tentativa de conectar-se à Internet demora demais, o serviço de quarto pode levar horas, o serviço de	negativo

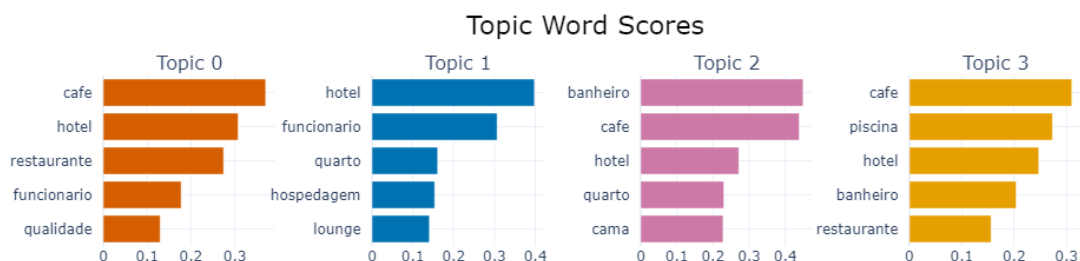
	despertar são feitos no fuso horário chinês ou é esquecido, não há check-in antecipado ou check-out tardio, e a última vez que alguém sorriu na recepção (o que incomum para os brasileiros) foi um ano atrás. Não recomendado.	
3	Passei um final de semana no royal tulip. Primeiro pegamos uma fila enorme para fazer p check in, quartos limpos porém um pouco antigo. A água do frigobar e 300 ml e custa um absurdo, na piscina só tem cerveja em lata, que também é um absurdo de caro. Espera muito mais. Atendimento muito demorado também.	negativo
3	Se quiser ter noção de onde estão os pontos turísticos, é só dar uma passada na cobertura, onde fica a piscina . Mas não se esqueça de levar escova de dente e pasta. Se não voce tera que fazer como eu que teve que comprar nas máquinas de atendimento automático.	negativo

10.6. HOTEL GRAND HYATT DE SÃO PAULO

O método foi aplicado ao hotel Grand Hyatt de São Paulo, identificado no conjunto de dados de hotéis brasileiros obtidos da plataforma TripAdvisor, especificamente filtrando as avaliações relacionadas ao hotel em questão. Embora o conjunto de dados não incluía uma coluna identificando o hotel associado a cada comentário, foi possível identificar as avaliações relacionadas ao Grand Hyatt por meio de um padrão no arquivo CSV. Os comentários estão organizados cronologicamente no arquivo, o que permite estimar a qual hotel pertencem com base na sua posição. Dessa forma, foram identificados os índices associados explicitamente ao hotel Grand Hyatt para obter todos os comentários correspondentes, totalizando 2.623 avaliações desde o primeiro comentário mencionando o nome do hotel até o último.

Para que fosse possível ter uma demonstração visual das palavras mais mencionadas nos comentários das avaliações foi gerada uma nuvem com tais palavras, onde é possível visualizar que palavras como “hotel”, “serviço”, “manhã” e “atendimento” foram frequentemente escritas.

Prosseguindo com a modelagem de tópicos, foram identificados 4 tópicos, sendo estes os escolhidos para análise.



A seguir, é apresentada a aplicação completa do método para o tópico 0, incluindo as frases descritivas associadas. Também são mostrados os resultados da análise de sentimentos, com a quantidade de comentários classificados como “positivos”, “negativos” e “neutros”. Além disso, são listadas as entidades nomeadas encontradas e suas frequências. Por fim, são exibidos os três primeiros comentários relacionados ao tópico, categorizados por sentimento.

Tópico	Palavras do Tópico	Frequência
--------	--------------------	------------

0	[('cafe', 0.3705641), ('hotel', 0.3076537), ('restaurante', 0.27457952), ('funcionario', 0.1776864), ('qualidade', 0.12977296), ('equipe', 0.11755521), ('hospedagem', 0.10316763), ('banheiro', 0.08565643), ('evento', 0.061486628), ('bar', 0.057706308)]	900
---	--	-----

Modelo de Linguagem	Tópico	Frase
ChatGPT	0	O tópico trata sobre a experiência gastronômica no Grand Hyatt, abordando aspectos como café, restaurante e qualidade dos serviços.
ChatGPT	0	Este tópico discute a experiência de café e restaurante no Grand Hyatt, destacando também a qualidade dos serviços.
ChatGPT	0	Trata-se da análise da experiência gastronômica no Grand Hyatt, com foco em café, restaurante e a qualidade dos serviços oferecidos.
Gemini	0	O tópico trata sobre a experiência de hospedagem no Grand Hyatt, com foco em café da manhã, qualidade do serviço, equipe atenciosa e acomodações confortáveis.
Gemini	0	O tópico aborda os destaques do Grand Hyatt, incluindo café da manhã delicioso, atendimento de qualidade por parte da equipe, quartos bem equipados e serviços completos para uma estadia agradável.
Gemini	0	O tópico foca em desfrutar da hospitalidade do Grand Hyatt,

		mencionando café da manhã de qualidade, equipe atenciosa, quartos confortáveis e serviços que atendem às necessidades dos hóspedes para uma experiência memorável.
--	--	--

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
0	812	23	65

Tópico	Entidades frequentes	Frequência
0	sao, paulo, hyatt, morumbi, grand, hy, japon, brasil, franc, paulista	225, 214, 42, 36, 31, 28, 25, 24, 13, 12

Tópico	Comentário	Sentimento
0	Ótima opção para um brunch de domingo à tarde. Super bem servido, ótimo custo benefício, bom atendimento, recomendo porque vale a pena! Opções de saladas, pratos quentes, churrasco (com ótimas carnes), sobremesa e chopp a vontade. Para completar, uma música ao vivo com uma cantora Maravilhosa!	positivo
0	Se está procurando uma opção de estadia em SPaulo esta é a opção a ser escolhida. Quarto extremamente confortável, atendimento da equipe excelente , ótimo café da manhã, chuveiro maravilhoso, tudo vale à pena neste hotel! E apesar de toda a imponência tudo é simples.	positivo
0	Estivemos nesse hotel passando o carnaval, com toda a família Grossi e foi maravilhoso! Destaque para o Brunch e o lounge executivo que estavam excederá! Equipe muito atenciosa, em especial a Ninive da	positivo

	recepção.	
0	Serviço profissional com funcionários discretos e atenciosos. Apartamento com instalações confortáveis e limpas. Ar condicionado poderia funcionar melhor pois só ventilava. Serviço de quarto rápido e café da manhã diversificado com opções de frutas, pães doces e salgados tudo muito fresco.	neutro
0	Tudo que se refere ao hotel, digo, ao prédio mesmo, funciona muito bem, recepção, café da manhã, manobristas. Entretanto o jantar di Dia dos Namorados do restaurante Eau deixou muito a desejar, tanto em serviço como em sabor. Felizmente se redimiram no Brunch do Domingo.	neutro
0	Bom atendimento , mas o carpete do quarto surrado, o colchão para um cinco estrelas deixa a desejar e apesar de existir o menu travesseiros , um que uso de forma firme então tinha, ou seja falta refino na finalização de um excelente hotel	neutro
0	É contada uma taxa de R\$ 20 para entrega no quarto. Foram a sopa e o suco de laranja mais caros (total R\$ 81,40) que já paguei na vida. Para completar a sopa (minestrone) estava rala e sem graça. Uma péssima relação custo-benefício!	negativo
0	Quando nos hospedamos por conta de um congresso me surpreendi com uma recepcionista que não me lembro o nome que zombou de minha cara quando perguntei se havia duchas higiênicas nas suítes. Comuniquei com o gerente do dia que para corrigir tal constrangimento que passei, me deu acesso exclusivo de elevador. Como se isso afagasse meu descontentamento. Como formador de opiniões sinceramente tenho ainda meus receios apesar de isto	negativo

	ter acontecido há mais tempo do que o informado; mesmo porque já estive em outros hotéis da rede sem este tipo de problema. Ainda bem que os check in são reservados caso contrário o meu constrangimento poderia ser maior. De resto o que se tem a dizer do Grand Hyatt são só elogios tirando este grande dissabor.	
0	Na data de nosso check in, não sei o que houve! Chegamos antes do almoço, lá pelas 11:00 h. Porém, apenas liberaram nosso apartamento às 18:00 h! Realmente não entendemos o que aconteceu naquela ocasião! Ficou a desejar.	negativo

A seguir, é apresentada a aplicação completa do método para o tópico 1, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
1	[('hotel', 0.3979951), ('funcionario', 0.30667144), ('quarto', 0.16100049), ('hospedagem', 0.15362684), ('lounge', 0.14075863), ('equipe', 0.13596147), ('localizacao', 0.12593475), ('lugar', 0.11549643), ('estadia', 0.107271), ('restaurante', 0.08044033)]	547

Modelo de Linguagem	Tópico	Frase
ChatGPT	1	O tópico trata sobre a experiência geral no Grand Hyatt, com destaque para o hotel, funcionários e quartos.
ChatGPT	1	Este tópico aborda a experiência e serviços oferecidos no Grand Hyatt, com foco no hotel, funcionários e quartos.

ChatGPT	1	Trata-se da análise da estadia no Grand Hyatt, com ênfase no hotel, funcionários e acomodações.
Gemini	1	O tópico trata sobre a hospedagem no Grand Hyatt, com foco em quartos confortáveis, equipe atenciosa, localização privilegiada e lounge para relaxar, tornando a estadia memorável.
Gemini	1	O tópico aborda os diferenciais do Grand Hyatt, incluindo quartos bem equipados, atendimento de qualidade por parte da equipe, localização estratégica e espaço aconchegante para um descanso completo.
Gemini	1	O tópico foca em desfrutar de uma excelente estadia no Grand Hyatt, mencionando quartos confortáveis, equipe atenciosa, localização vantajosa e lounge para momentos de lazer, proporcionando uma experiência única aos hóspedes.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
1	505	18	24

Tópico	Entidades frequentes	Frequência
1	sao, paulo, hyatt, morumbi, grand, hy, brasil, paulista, globo, pa	144, 134, 31, 18, 16, 15, 11, 9, 4, 4

Tópico	Comentário	Sentimento
---------------	-------------------	-------------------

1	Foi nosso aniversário de casamento. Passamos de sábado para domingo. Gostamos muito. O atendimento é excelente. Adoramos o lounge Chandon, pena que tinha poucas pessoas. Talvez tenha que ser mais divulgado. Amei o som na piscina .	positivo
1	Tudo ocorreu perfeitamente, conforme era de se esperar de um hotel da prestigiada rede Hyatt. Alto nível de conforto, excelência e profissionalismo do staff. Sugiro apenas que passem a oferecer um turndown service.	positivo
1	excelente experiência a família. hotel com infraestrutura muito boa. Ótima localização, com shoppings, cinemas, teatros e outras facilidades muito próximas. voltaria com toda certeza. xv xv vx vx xv xv	positivo
1	hotel já foi mais requintado , atento aos detalhes, caiu muito em relação a minha estada anterior , principalmente nas refeições, e é inadmissível num hotel deste porte ter louças rachadas ou faltando pedaços.	neutro
1	O valor não condiz com o que oferece, hotel igual a muitos outros e o pacote dos dias do namorado não tem um diferencial. Não estou dizendo que o hotel seja ruim, pelo contrário é bom, mas pelo valor da diária com certeza encontramos muitos outros hotéis com melhor custo benefício. O quarto é grande, mas não há opções de frigobar	neutro
1	ótimo hotel em todos os sentidos, porém não posso aceitar passar duas noites de verão sem ar condicionado, apesar de, por três vezes, ter recebido da recepção a promessa de solucionar o problema imediatamente. Isso provocou duas noites mal dormidas.	neutro

1	Uma vez a cada dois meses eu e meu marido vamos a passeio para SP, e sempre nos hospedamos ou no Renaissance ou no Hilton, é sempre foi sensacional, sempre! Ocorre que como ouvimos falar do Grand Hyatt, resolvemos nos hospedar lá por ter o mesmo estilo com club igual aos demais. Ocorre que já no contato telefônico fui muito mal tratada, primeiro me atendeu a Vitória que falou que não ia esclarecer minhas dúvidas que faltava 20 minutos para a central de reserva abrir para eu ligar depois, mesmo eu explicando que só queria saber como funcionava o club, não queria saber valores. Liguei uma hora depois a mesma me passou para a central de reserva o qual atendeu um homem grosseiro sem paciência algum, e além de tudo irônico. Estava ansiosa para passar meu aniversário de casamento nesse hotel. Pois é perderam cliente, pois o tratamento que dá nome ao hotel começa na reserva. Decepcionante	negativo
1	Podem melhorar e muito, seja com relação a atendimento, quarto e hotel. Deixou muito a desejar!!! As toalhas estavam esgarçadas, faltava itens no banheiro. Frigobar desabastecido, muita incidência de pernilongo nas dependências do hotel.	negativo
1	Fiz todos os preparativos necessários para uma noite de pedido de casamento, liguei antes do check-in para adiantar todo o processo mas o hotel estava despreparado, o pacote romance estava incorreto. O quarto com a porta quebrada e empenada. Cortina suja, o champanhe não estava no quarto, lençois e edredon também não... Pecaram em vários detalhes nesta data importante e por este motivo não pretendo voltar. Certeza que um hotel de menor porte daria mais atenção para esta data. Ainda bem que não foi lua de	negativo

	mel, pois estaria bem mais frustrados. Espero que melhorem.	
--	---	--

A seguir, é apresentada a aplicação completa do método para o tópico 2, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
2	[('banheiro', 0.44947478), ('cafe', 0.43861878), ('hotel', 0.27158695), ('quarto', 0.22998606), ('cama', 0.22804582), ('lounge', 0.19150598), ('piscina', 0.1665397), ('estadia', 0.16003507), ('acomodacao', 0.14699297), ('restaurante', 0.14049123)]	337

Modelo de Linguagem	Tópico	Frase
ChatGPT	2	O tópico trata sobre as comodidades do banheiro e café no Grand Hyatt, seguido pela experiência geral no hotel.
ChatGPT	2	Este tópico aborda a qualidade e comodidades dos banheiros e cafés no Grand Hyatt, destacando a experiência no hotel.
ChatGPT	2	Trata-se da análise das instalações de banheiro e café no Grand Hyatt, com foco na experiência geral de hospedagem..
Gemini	2	O tópico trata sobre o conforto e relaxamento no Grand Hyatt, com foco em banheiros impecáveis, café da manhã delicioso, quartos aconchegantes, camas confortáveis e área de piscina para uma estadia relaxante.

Gemini	2	O tópico aborda as comodidades do Grand Hyatt, incluindo banheiros completos, café da manhã de qualidade, quartos bem equipados, camas macias, área de lazer com piscina e serviços completos para atender às necessidades dos hóspedes.
Gemini	2	O tópico foca em desfrutar de uma experiência relaxante no Grand Hyatt, mencionando banheiros modernos, café da manhã saboroso, quartos convidativos, camas confortáveis e área de piscina para um descanso completo, proporcionando uma ótima estadia aos hóspedes.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
2	321	6	10

Tópico	Entidades frequentes	Frequência
2	sao, paulo, morumbi, hyatt, grand, brasil, hy, globo, americano, rede	61, 58, 12, 9, 8, 7, 6, 4, 4, 3

Tópico	Comentário	Sentimento
2	Hotel de excelente qualidade. Quarto amplo e muito confortável , banheiro perfeito com amenities de excelente qualidade. A cama e os travesseiros extremamente confortáveis , o atendimento de primeira categoria. Sem duvida uma excelente hospedagem !	positivo
2	piscina com espreguiçadeiras confortáveis, quartos acolhedores , bem decorados , ótima ducha. A Ponte	positivo

	estaiada vista da janela é mais um flash na lembrança de um final de semana delicioso . Estamos sempre por lá para um final de semana na Capital .	
2	O Hotel é muito bonito e o atendimento é muito prestativo, mas nos quartos já é possível notar a sua idade. A limpeza do banheiro é razoável e alguns móveis já não tem um bom estado de conservação. Recomendo pesquisar bastante o preço das diárias antes de reservar pois elas podem oscilar bastante! Possui um ótimo café da manhã com muitas opções.	positivo
2	As refeições de piscina deveriam ter mais opções para o que é um hotel de luxo. Chopp não estava gelado e nos disseram que estava. No geral tudo ok Os banheiros da piscina precisam de uma reforma. Os restaurantes são excelentes.	neutro
2	Achei a cama de casal muito pequena, não é condizente com o que espero de um hotel do nível. O ar condicionado do quarto também não funciona muito bem, o que tornou a noite de sono bem desagradável. O hotel é lindo, a vista do quarto espetacular (principalmente se vc fica nos quartos mais altos) e o banheiro é bem luxuoso. O atendimento é muito bom. Ah, a localização do hotel também não é das melhores.	neutro
2	Quarto confortável, porém, nao serve para quem gosta de escuridão total para dormir. Desde cedo ja entra luz do dia pelas frestas das cortinas. Se vc procura alguma privacidade opte por quartos em andares altos. Caso contrário, vc vai abrir as cortinas e se deparar com o prédio da Rede Globo. Produtos do banheiro excelentes, porém, pouca quantidade.	neutro
2	Fiquei no quarto 2221 e o quarto apresentou diversos problemas, que apesar se terem sido acionados para a	negativo

	solução do mesmo, tiveram de ser acionados novamente 2 e ate 3 vezes. A cortina do quarto não subia (acionei 3 vezes), ar condicionado nao gelava (acionei 2 vezes), chuveiro sem pressão e banheira sem demorava horas para encher.	
2	Apesar de ser um "hotel para negócios" creio que deveriam ter mais atenção com os detalhes, como por exemplo, deixar roupões de banho no banheiro e nunca vi um frigobar tão "pobre": não havia um petisco salgado sequer, apenas uma embalagem de balinhas azedas e nem opções de refrigerantes diet. Estivemos anteriormente hospedados, em junho de 2015, e emendamos o fds do dia dos namorados... estava melhor...	negativo
2	a cama era pessima, com um buraco no meio, o que fez eu ficar com muita dor na coluna. os travesseiros tambem, muiot desconfortaveis, e a opcao de troca, pior ainda. o banheiro era bom, mas sem duchinha higienica, apenas bide; e o secador de cabelos muito fraco.	negativo

A seguir, é apresentada a aplicação completa do método para o tópico 3, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
3	[('cafe', 0.31035864), ('piscina', 0.2733199), ('hotel', 0.24709561), ('banheiro', 0.2039112), ('restaurante', 0.15608785), ('almoco', 0.12658373), ('hospedagem', 0.11094935), ('quarto', 0.084445454), ('lounge', 0.081025265), ('bar', 0.07985477)]	181

Modelo de Linguagem	Tópico	Frase
ChatGPT	3	O tópico trata sobre a experiência com café e piscina no Grand Hyatt, seguido pela descrição geral do hotel.
ChatGPT	3	Este tópico aborda a qualidade do café e as comodidades da piscina no Grand Hyatt, com ênfase na experiência de hospedagem.
ChatGPT	3	Trata-se da análise da oferta de café e piscina no Grand Hyatt, destacando a experiência geral no hotel.
Gemini	3	O tópico trata sobre momentos de lazer no Grand Hyatt, com foco em piscina refrescante, café da manhã delicioso e restaurante com opções para almoço, proporcionando uma estadia relaxante e agradável.
Gemini	3	O tópico aborda as opções de lazer e gastronomia do Grand Hyatt, incluindo piscina para se refrescar, café da manhã de qualidade, restaurante com cardápio variado para o almoço e serviços completos para atender às necessidades dos hóspedes.
Gemini	3	O tópico foca em desfrutar das comodidades do Grand Hyatt, mencionando piscina para um mergulho refrescante, café da manhã saboroso, restaurante com opções para o almoço e serviços que proporcionam uma experiência completa aos hóspedes.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
3	170	4	7

Tópico	Entidades frequentes	Frequência
3	sao, paulo, japon, morumbi, grand, hyatt, frances, franc, olimpia, hy	53, 50, 12, 9, 8, 6, 4, 3, 3, 3

Tópico	Comentário	Sentimento
3	Um Hotel elegante, charmoso, aconchegante, com um Staff de primeira, com pessoas educadas, delicadas, gentis. Estive hospedada entre os dias 9 e 13/02, durante o carnaval 2018. Os quartos são amplos, bem limpos e confortáveis. O café da manhã sempre bem servido e com uma grande variedade. Utilizei o serviço de Spa que o Hotel dispõe, e também adorei, indico. Havendo disponibilidade de tempo vale a pena dar uma passadinha por lá, precisa agendar horário com antecedência, pois é bem disputado. O Hotel ainda dispõe de uma academia super bem equipada, piscina aquecida no Spa e natural ao ar livre na área externa, ambas são maravilhosas... a externa é super concorrida, necessário já reservar a espreguiçadeira antes de tomar o café da manhã, embora haja uma boa quantidade delas é de fato concorrida a ocupação das mesmas... Contando os dias para retornar...	positivo
3	Hotel bem localizado, tem um volume alto de quartos, mas infraestrutura para atender bem todos. Os quartos são espaçosos e a equipe super atenciosa. O café da manhã é completo e se tiver um dia agradável pode ser na beira da piscina . Ah, a piscina é super bacana, mas	positivo

	pega sol só no período da tarde.	
3	Ótimo Hotel da rede Hyatt, atendimento padrão Hyatt, principalmente no Club Hyatt! Vale a pena!! Bem localizado para hospedagem de finais de semana, silencioso (longe da rota de aviões), quartos e banheiros amplos, bem completos e super limpos. Ótima piscina, bem como o atendimento no bar desta. Excelente café da manhã, muitas opções saudáveis e muita variedade, agrada a todos! Ótima vista da cidade e da ponte estaiada!! Recomendo e volto com certeza!!	positivo
3	O Hotel é lindo, conta como uma equipe muito cordial mas poderia contar com algumas melhorias para chegar a perfeição esperada em um hotel desse padrão. Ar condicionado nas áreas comuns são ligados em uma temperatura muito baixa, mas o ar dos quartos não funcionou direito. Chuveiro sem pressão e água que oscilava a temperatura durante o banho (hora gelado, hora escaldante). Serviço de Jantar bem demora, esperamos mais de 40 minutos por 2 hamburguers e ainda for serviço sem a batata (que chegou apenas 10 minutos depois). Batata frita era congelada, que ao meu ver, é inaceitável para o padrão e valor que se paga. Eu sou intolerante e não tinha opções sem glúten ou leite no café da manhã. Quando solicitei ao garçom ele me serviu um bolo industrializado. Muitos hotéis, inclusive de padrão inferior, já servem opções de bolos, pães, quaijos e leites alternativos.	neutro
3	Em termos gerais, estão de parabéns. Entretanto, pude observar no período em que estive hospedado com minha família alguns pontos negativos: Na suíte em que ficamos hospedados, no vigésimo segundo andar, notamos ao chegarmos, que haviam fezes no vaso sanitário - o que nos causou uma má impressão com	neutro

	relação a higiene da suíte, além disso, minha esposa notou diversos pêlos ao redor da banheira e a mesma esta "lisa / engordurada"; Não haviam acomodações suficientes para os hóspedes na piscina localizada no térreo e apesar de termos utilizado a piscina coberta no segundo andar, observamos que diversas pessoas ficaram mal acomodadas ou desistiram de utilizar a piscina ; O almoço somente começou a ser servido na área de alimentação após às 15:00 horas devido ao café da manhã extendido. Creio que o café da manhã extendido - opcional, deveria ser servido nos quartos, após o horário normal, somente para aqueles que deste serviço precisassem e não impor esta condição a todos, principalmente pelo fato de que os restaurantes adjacentes não tinham capacidade de atendimento a todos os hóspedes. Pro fim, no ato da reserva, não fomos informados que os restaurantes adjacentes somente trabalhariam com pré-reserva no dia primeiro. Caso tivéssemos sido informados das alterações no funcionamento do restaurante do hotel devido ao procedimento de café da manhã extendido, teríamos feito a pré-reserva, pois temos um filho de quatro anos de idade que foi o principal lesado com toda esta confusão.	
3	A piscina externa poderia ser climatizada. Mesmo com o dia ensolarado a água estava um gelo, inviável permanecer .	neutro
3	Uma diária romântica . Pedi fotos no quarto que foi uma ideia genial . Comprei a champa que ela mais gosta . Estava gelada . Quarto com vista boa . Acabou aí a parte legal . Piscina descoberta em reforma que não fui avisado nem na reserva , nem no check in (mesmo perguntando qual a temperatura da piscina onde fui	negativo

	informado que não era aquecida)Café da manhã sem graça com frutas com casca cortadas de qualquer jeito , todos os funcionários de ma vontade , principalmente a hostess aliene Garçons não sabiam o que era ovo Benedictine , quais tipos de espresso tinham Paes sem graca e frios também Padrao Ibis com diária de 1500 reais Coquetel no andar club com Chandon nacional e vinhos chilenos de 30 reais Comida barata e sem graca mas a garçonete era muito atenciosa Quarto precisa de Reforma urgenteCama só tem lençol e edredom . Não tem sobre lençol Ou voce morre de frio ou de calor Pedi lençol extra e travesseiros as 23 h que chegaram meia noite Estacionamento 53 reais Piscina aquecida não pode tomar bebida alcoólica e você divide a sua privacidade com executivos que vem fazer spa avulso e ficam olhando sua mulher Enfim pelo que cobra ... pode melhorar muuuuito Pelos concorrentes em SP ele terá que mudar muito para eu mudar minha avaliação.	
3	Ficamos de 27 a 29/01/17 no apto 1023. Esperava mais do hotel. A arrumação do apto no sábado só ocorreu às 15:00 horas e durou quase 40 min, só uma camareira. Qnd fui usar as piscina s estavam muito cheias, sem mesas vagas.	negativo
3	Esperava mais do hotel por ser considerado um dos melhores de Sao Paulo. A estrutura é excelente - lindos quartos, duas piscina s, boa comida. Mas o ar condicionado é MUITO fraco. As noites foram muito quentes. Em uma das noites teve uma manutencao eletrica, onde propositalmente se desligou toda eletricidade por 4-5 horas - em uma das noites mais quentes do ano. Foi insuportavelmente incomodo. A academia apesar de ter um excelente tamanho para uma academia de hotel, aceita não hospedes (parece	negativo

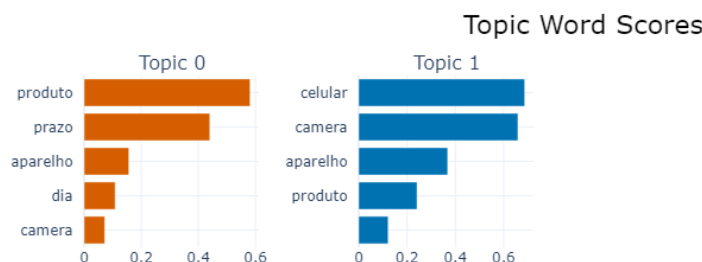


Figura 31 - Gráfico de barras dos 2 tópicos mais frequentes do conjunto de dados do smartphone da Motorola. Fonte: A autora.

A seguir, é apresentada a aplicação completa do método para o tópico 0, incluindo as frases descritivas associadas. Também são mostrados os resultados da análise de sentimentos, com a quantidade de comentários classificados como “positivos”, “negativos” e “neutros”. Além disso, são listadas as entidades nomeadas encontradas e suas frequências. Por fim, são exibidos os três primeiros comentários relacionados ao tópico, categorizados por sentimento.

Tópico	Palavras do Tópico	Frequência
0	[('produto', 0.5815523), ('prazo', 0.4405504), ('aparelho', 0.1568595), ('dia', 0.108938314), ('camera', 0.07238525), ('celular', 0.06567275), ('', 0.03450574), ('', 0.03450574), ('', 0.03450574), ('', 0.03450574)]	513

Modelo de Linguagem	Tópico	Frase
ChatGPT	0	O tópico trata sobre a análise do produto da Motorola, com foco no prazo de entrega e nas características do aparelho.
ChatGPT	0	Este tópico aborda a análise geral do produto da Motorola, com destaque para o prazo de entrega e as especificações do aparelho.

ChatGPT	0	Trata-se da análise do produto da Motorola, com ênfase no tempo de entrega e nas características do dispositivo.
Gemini	0	O tópico trata sobre o lançamento de um novo produto da Motorola, com foco em prazos de entrega, características do aparelho e data de lançamento.
Gemini	0	O tópico aborda as informações sobre um novo celular da Motorola, incluindo data de lançamento, funcionalidades do aparelho e detalhes do produto.
Gemini	0	O tópico foca em desvendar o novo smartphone da Motorola, mencionando prazos, características e data de lançamento, despertando a curiosidade dos consumidores.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
0	438	20	55

Tópico	Entidades frequentes	Frequência
0	motorola, americanas, americana, moto, g5s, lenovo, plus, samsung, g5, android	27, 16, 13, 11, 10, 6, 4, 2, 2, 2

Tópico	Comentário	Sentimento
0	Produto excelente, ótimas funções, rápido, visual muito agradável.	positivo
0	Excelente produto, belo design, fácil de manusear. Super satisfeita.	positivo

0	Satisfaz todas as expectativas, ótimo aparelho, recomendo com certeza.	positivo
0	Produto bom, mas fiquei decepcionado com o volume do aparelho, muito, muito baixo, que pena.	neutro
0	Poderia ser melhor se a câmara traseira não fosse avançada do alinhamento traseiro	neutro
0	Este é o terceiro Motorola da minha casa, por insistência do meu filho que gosta da navegação dele, mas todos os outros até agora, tiveram que ser trocados porque dão problema na bateria. O produto é bom, rápido com boa navegação, até que pare de carregar. Isso acontece com menos de um ano.	neutro
0	Comprei o produto e veio com defeito. A americaas.com não realizou a troca. Estou muito decepcionada, visto que o produto era para presente.	negativo
0	COMPRI ESTE APARELHO, MAIS ESTAVA COM DEFEITO, A BATERIA ESTAVA AQUECENDO MUITO, E DESISTIR DA COMPRA.SENDO QUE A EMPRESA JÁ REEMBOLSOU.	negativo
0	Desde o dia 03\01\2018 que o produto está a caminho do meu endereço e até agora não chegou e nenhuma atualização no site. Já enviei varios emails para a transportadora e nada de resposta.	negativo

A seguir, é apresentada a aplicação completa do método para o tópico 1, incluindo todas as informações fornecidas assim como no tópico anterior.

Tópico	Palavras do Tópico	Frequência
1	[('celular', 0.68748057), ('camera', 0.6593984), ('aparelho', 0.36857295), ('produto', 0.2416237), ('', 0.122100055),	279

	('', 0.122100055), ('', 0.122100055), ('', 0.122100055), ('', 0.122100055), ('', 0.122100055)]	
--	--	--

Modelo de Linguagem	Tópico	Frase
ChatGPT	1	O tópico trata sobre a análise do smartphone da Motorola, com foco na qualidade da câmera e no desempenho do aparelho.
ChatGPT	1	Este tópico aborda a análise detalhada do celular da Motorola, destacando a excelência da câmera e o desempenho do aparelho.
ChatGPT	1	Trata-se da avaliação do smartphone da Motorola, com ênfase na câmera de alta qualidade e no funcionamento geral do dispositivo.
Gemini	1	O tópico trata sobre a qualidade das câmeras de um celular da Motorola, com foco em recursos inovadores, desempenho fotográfico e características do aparelho.
Gemini	1	O tópico aborda as especificações de um smartphone da Motorola, incluindo sistema de câmeras de alta resolução, funcionalidades do aparelho e detalhes do produto.
Gemini	1	O tópico foca em capturar fotos incríveis com um celular da Motorola, mencionando recursos avançados das câmeras, qualidade de imagem e experiência fotográfica superior.

Tópico	Comentários positivos	Comentários neutros	Comentários negativos
1	228	20	31

Tópico	Entidades frequentes	Frequência
1	motorola, moto, g5s, android, americanas, asus, lenovo, ios, google, gmail	20, 7, 5, 5, 5, 4, 2, 2, 2, 1

Tópico	Comentário	Sentimento
1	Produto excelente qualidade boa câmera desenvolvimento do Android com rapidez	positivo
1	Nada mal, tempo de entrega antes do prazo. Sobre o Celular ótimo atendeu muito bem ao que eu esperava. Bateria em média 1 dia e meio +/-.	positivo
1	Gostei do pedido e prazo de entrega. Celular muito bom para o preço pago, ótima promoção	positivo
1	O aparelho é bom, só que tem horas que o touch trava. Referente a câmera já vi melhores por ter 12 MP acho que a imagem deveria ser um pouco melhor.	neutro
1	Bom produto e bonito. Mas as vezes trava. E fica difícil de voltar as funções. Tem dificuldade de conexão com a internet. A bateria carrega rápido e dura bem.	neutro
1	Existe os pontos positivos e negativos... O celular tem ótimo funcionamento, porém não sabia que a bateria não era removível. Ele esquenta não sei se é normal. Não gostei do modo da colocação do chip, achei bem frágil.	neutro
1	Boa tarde. Estou insatisfeita com o aparelho, o mesmo está apresentando problemas na câmera frontal, na	negativo

	verdade a câmera está travada e não há o que fazer. Quero trocar o produto.	
1	transportadora foi infeliz desviaram o meu produto, estou muito triste com as lojas americanas não me dao uma resposta só enrolam falaram que iam retorna, retornou o carai, nem as lojas fisicas ajudam pessimo atendimento pra min isso é uma humilhação da transportadora um celular novo nao cai do céu	negativo
1	ão é avisado para quem compra que o celular só permite expansão de memória se for optado para o uso somente de um chip de telefone. Se usar 2 chips de para o telefone não pode usar cartão de memória. Não recomendo a compra.	negativo

11. AVALIAÇÃO TÉCNICA DO MÉTODO E SEUS RESULTADOS

Como foi apresentado na seção anterior, o método desenvolvido abrange a modelagem de tópicos, análise de sentimentos e extração de entidades nomeadas, aplicado a comentários de avaliações de produtos e serviços em diversas plataformas. Essas plataformas incluem, em ordem cronológica, o restaurante Guacamole, plataformas de comércio eletrônico como Olist, Buscapé e B2W Digital, hotéis brasileiros e, por fim, o hotel Grand Hyatt em São Paulo, além de um produto específico, o Smartphone da Motorola.

Seguindo os processos realizados, as seguintes conclusões foram assumidas:

- **Modelagem de Tópicos:** Na modelagem de tópicos observamos resultados interessantes, especialmente quando os tópicos abordam questões que vão além das avaliações usuais de produtos e serviços, como prazos de entrega, atendimento e qualidade do objeto comprado. Por exemplo, ao aplicar o método nas avaliações do Buscapé e B2W, foram identificados assuntos predominantes relacionados a perfumes e carros vendidos na plataforma, e também a livros e produtos que possivelmente foram presentes para membros da família dos compradores. Dessa forma, é possível notar a diversidade e amplitude das

avaliações especialmente presentes nesses conjuntos de dados. Logo, a modelagem de tópicos demonstrou utilidade conforme o esperado para o método, revelando resultados interessantes que transcendem as avaliações usuais de produtos e serviços.

- **Geração de Frases Descritivas sobre os Tópicos:** Na geração de frases descritivas sobre os tópicos usando modelos de linguagem como o ChatGPT e o Gemini, ficou evidente que cada um possui suas próprias características distintas. Enquanto o ChatGPT conseguiu descrever os tópicos de maneira mais concisa, suas descrições muitas vezes careciam de detalhes, utilizando frases que nem sempre transmitiam com coerência o assunto abordado, como observado na primeira frase gerada pelo ChatGPT no tópico de número 1 do conjunto de dados da B2W Digital onde o mesmo concluiu que aspectos das funcionalidades seriam destacados. Por outro lado, o Gemini demonstrou habilidade superior em criar frases mais elaboradas, porém ocasionalmente exageradas, tentando construir uma narrativa ou contexto emotivo que, por vezes, se distanciava das informações originais do tópico. No geral, o Gemini se destacou quando suas frases mais elaboradas foram descartadas, apresentando descrições mais coerentes e descritivas. Portanto, o Gemini foi escolhido devido à sua capacidade em gerar frases mais consistentes e informativas, ressaltando a importância da geração de tais frases para a compreensão dos tópicos analisados.
- **Análise de Sentimentos:** Na análise de sentimentos, notamos que os resultados apontam predominantemente para classificações positivas dos produtos e serviços avaliados. Isso sugere uma alta satisfação entre os consumidores em todos os casos analisados. A análise de sentimentos, portanto, confirmou sua utilidade esperada para o método, proporcionando insights valiosos sobre as percepções e emoções dos clientes em relação aos produtos e serviços em questão. Além de destacar os aspectos que agradaram aos consumidores, também revelou áreas de insatisfação, possibilitando uma compreensão mais completa da experiência do cliente.
- **Extração de Entidades Nomeadas:** Na extração de entidades nomeadas, deparamo-nos com resultados interessantes, especialmente nas avaliações do Buscapé onde foram identificadas marcas de perfumes e carros. No entanto, identificamos também alguns casos curiosos, como nas avaliações da Olist, onde

adjetivos foram erroneamente considerados entidades. Além disso, destacaram-se situações em que nomes de pessoas e marcas de produtos foram frequentemente identificados como entidades. Portanto, a extração de entidades nomeadas demonstrou ser útil para o método, especialmente em análises que abrangem uma variedade de temas, indo além de produtos ou serviços específicos.

- **Análise dos Comentários:** Na análise dos comentários, mesmo com apenas três exemplos de cada variação de sentimento para cada tópico, podemos observar que as relações com os tópicos são coerentes. No entanto, uma apresentação dinâmica que permita uma paginação dos comentários, possibilitando a visualização de todos os disponíveis caso desejado, seria mais útil para uma compreensão mais abrangente da visão dos consumidores sobre os temas identificados. Em relação aos sentimentos, notamos que o modelo treinado é capaz de identificar bem os sentimentos positivos e negativos, mas apresenta certa dificuldade em classificar os comentários como neutros, às vezes rotulando comentários claramente positivos como neutros. Isso pode ser observado nas avaliações classificadas como neutras no tópico de número 0 do conjunto de dados do restaurante Guacamole.

Inicialmente, a aplicação do método foi planejada para ser executada nos conjuntos de dados do restaurante Guacamole e nas plataformas de comércio eletrônico da Olist, Buscapé e B2W. No entanto, ao analisar os resultados desses casos, identificamos um problema fundamental: a dificuldade em relacionar um produto específico a um tópico ou comentário, uma vez que todos os comentários estão implicitamente relacionados a um produto, sem mencionar explicitamente seu nome. Diante dessa questão, decidimos expandir a aplicação do método para o conjunto de dados de hotéis brasileiros para verificar se esse padrão se mantinha ou se poderia ser aplicado de forma mais eficaz e, como esperávamos, o resultado foi bem semelhante. Assim, optamos por realizar a filtragem nos dados a fim de aplicar aos dados do hotel Grand Hyatt de São Paulo e no produto Smartphone da Motorola para uma análise mais aprofundada, especificando os domínios.

Dessa forma, foi possível concluir que o método produz resultados mais claros e consistentes quando o conteúdo das avaliações é especificado, permitindo uma análise menos generalizada e mais precisa.

Em termos gerais, o método demonstrou eficácia na identificação das principais preocupações e satisfações dos clientes em relação aos produtos e serviços, ao mesmo tempo em que identifica os sentimentos associados a eles, proporcionando insights detalhados por meio dos comentários e das entidades relacionadas.

12. AVALIAÇÃO DA UTILIDADE DO MÉTODO POR ESPECIALISTAS

Para avaliação do método, foram entrevistados dois especialistas em análise de dados com ampla experiência na área. O método foi demonstrado e explicado, juntamente com os resultados obtidos ao aplicá-lo ao conjunto de dados das avaliações do Smartphone da Motorola na plataforma Buscapé. As entrevistas foram realizadas com um especialista que possui vasta experiência como Gerente de Ciência de Dados e atualmente atua como Gerente Sênior de Engenharia de Software na OLX Brasil há dois anos. O mesmo concluiu seu Doutorado em Filosofia na Ciência da Computação na PUC-Rio em 2020 e teve seu primeiro contato com ciência de dados de forma profissional em 1999. Além disso, as entrevistas foram conduzidas com um outro especialista que tem uma extensa carreira como Analista de Dados e atualmente trabalha na TIM de telecomunicações como Diretor Executivo em Ciência de Dados há 2 anos. O mesmo completou seu Mestrado em Administração na PUC-Rio em 2007 e iniciou sua trajetória profissional em ciência de dados em 2008.

As entrevistas foram conduzidas de forma remota, onde foi exibida uma apresentação sucinta sobre a funcionalidade do método e como ele foi construído, disponibilizada na seção de Anexos. A entrevista com um especialista durou 24 minutos, enquanto a com o outro teve uma duração de 52 minutos. Ao final, foram feitas perguntas específicas sobre o método. Em alguns momentos da entrevista, reforcei detalhes sobre a construção do método, como a forma de obtenção dos dados, os algoritmos e modelos de linguagem utilizados, entre outros pontos.

Vale ressaltar que os especialistas não utilizaram o método diretamente, ou seja, não aplicaram o método em um conjunto de dados escolhido por eles. Durante as entrevistas, foram apenas apresentados os resultados obtidos a partir do conjunto de dados das avaliações do restaurante Guacamole.

- Como você avalia a facilidade de compreensão do método?

- Como você avalia a usabilidade do método?
- Como você avalia a utilidade (relevância, impacto) em saber os principais assuntos discutidos nas avaliações dos clientes de uma empresa?
- Como você avalia a utilidade (relevância, impacto) dos sentimentos extraídos pelo método?
- Como você avalia a utilidade (relevância, impacto) das entidades nomeadas extraídas pelo método?
- Em sua opinião, o método fornece insights úteis para melhorar a qualidade de produtos e serviços?
- Na sua avaliação, o método pode ser aprimorado de alguma forma? Se sim, de que forma?

De forma geral, ambos os especialistas consideraram o método fácil de compreender e sua aplicação satisfatória, destacando a boa estrutura do método. Além disso, foi observado que o método pode permitir relacionar avaliações de produtos e serviços com os perfis dos usuários em casos onde todas as informações da compra estão disponíveis, como o histórico de compras de um consumidor, possibilitando um outro caso de uso como, por exemplo, a exibição de informações sobre produtos bem avaliados que possam interessar a um usuário específico.

Ambos destacaram a importância da extração dos principais tópicos, sentimentos e entidades nomeadas para uma melhor compreensão dos consumidores e da visão da empresa sob a ótica do público. Com a crescente preocupação das empresas com a experiência do cliente (CX), eles concluíram que o método pode fornecer diversos insights valiosos para a melhoria da qualidade de produtos e serviços. Essa capacidade de identificar rapidamente insatisfações e principais problemas permite que a área de atendimento ao cliente ofereça suporte mais eficaz, pois já possui informações detalhadas sobre os pontos críticos destacados pelos usuários. Com isso, o método dá um suporte significativo para a área de atendimento ao cliente, facilitando a identificação e resolução de problemas.

Em particular, sobre a utilidade das entidades nomeadas, eles observaram que, embora em contextos específicos, como a análise de um produto, a identificação de entidades possa ser menos reveladora por ter um contexto bem definido para análise,

em análises mais gerais, essa informação é crucial para discriminar os diferentes assuntos discutidos. Saber quais entidades são mencionadas frequentemente é importante para entender sobre o que os clientes estão falando, além de fornecer um contexto mais rico sobre suas avaliações.

Em relação ao aprimoramento do método, ambos os especialistas destacaram alguns pontos importantes. Primeiramente, enfatizaram a importância da utilização de dados "Ground Truth" para garantir respostas mais confiáveis. Isso envolve ter um conjunto de dados de treinamento com uma curadoria rigorosa para assegurar que as avaliações sejam precisas, como nos casos em que a classificação por estrelas pode não corresponder aos comentários escritos.

Além disso, foi sugerida uma obtenção aprofundada de métricas para comprovar a eficácia das análises presentes no método. Isso é crucial para demonstrar a confiabilidade do método para uma empresa e testá-lo em produção em cenários reais já que, para implementar o método em produção em uma empresa, são necessárias mais evidências de que tudo está funcionando corretamente. Em um contexto de pesquisa e treinamento, o método está ótimo, mas é necessário identificar o que pode ser aprimorado para aplicá-lo em um ambiente real.

Outro ponto abordado foi a possibilidade de usar modelos de linguagem de grande porte (LLMs) para todo o método, visando melhorar a assertividade e a escalabilidade do método. Isso inclui o uso de métodos de otimização como o RAG (Retrieval-Augmented Generation), um processo de aperfeiçoamento da saída de um modelo de linguagem, para melhorar a aplicabilidade do método. Ao utilizar a análise de texto com comentários, nem sempre um comentário estará relacionado a um único tópico, havendo a possibilidade de serem multi-tópicos. Portanto, tornar o método escalável para o uso de LLMs é essencial, pois novos comentários e tópicos podem surgir ao longo do tempo, exigindo que o modelo seja revisado regularmente. O uso do TF-IDF exige o ajuste de muitos parâmetros, enquanto os LLMs podem proporcionar uma implementação mais eficiente e escalável. Assim, o método é útil, mas há formas mais avançadas de implementação que podem ser exploradas para otimizar os resultados.

Por fim, de forma sintética, eles consideraram que a facilidade de compreensão dos resultados fornecidos pelo método é elevada assim como o entendimento do seu uso. Foram igualmente avaliados positivamente a utilidade de conhecimento dos

principais assuntos, dos sentimentos e das entidades nomeadas extraídas do método, onde concluíram que o mesmo fornece resultados que podem ser bem utilizados para melhorar a experiência do cliente. Foi concluído que alguns ajustes são necessários de forma que deixe o método mais preciso e confiável para uso em cenários reais mas, em um geral, ele foi bem construído.

13. CONSIDERAÇÕES FINAIS

Esse trabalho teve como objetivo desenvolver e validar um método para avaliar a satisfação dos clientes em relação a serviços e produtos oferecidos no contexto do e-commerce brasileiro, empregando técnicas de mineração de texto ao abranger aspectos como modelagem de tópicos, análise de sentimentos e extração de entidades nomeadas. Através da aplicação deste método a diferentes conjuntos de dados, incluindo avaliações de restaurantes, plataformas de comércio eletrônico e hotéis, foi procurado fornecer insights prestativos para as empresas serem capazes de melhorar a qualidade de seus produtos e serviços, assim como a experiência de seus clientes.

O objetivo do trabalho foi alcançado, pois o método desenvolvido demonstrou ser eficaz em fornecer diversas informações valiosas em diferentes contextos. Com a aplicação das três técnicas (modelagem de tópicos, análise de sentimentos e extração de entidades nomeadas) foi possível obter insights sobre as percepções dos consumidores. Tais informações podem ser utilizadas para melhorar a compreensão das opiniões dos clientes, facilitando a tomada de decisões estratégicas para aprimorar produtos e serviços.

Durante o desenvolvimento do trabalho, algumas dificuldades significativas foram encontradas. Primeiramente, houve uma dificuldade em encontrar conjuntos de dados que fossem confiáveis, reais e com boa qualidade, sendo em língua portuguesa. Além disso, o ajuste dos parâmetros dos métodos do BERTopic para cada conjunto de dados foi um desafio, exigindo diversos testes para determinar quais configurações entregavam os melhores resultados. A busca por um modelo eficiente de extração de entidades nomeadas em português também foi complexa, devido à escassez de recursos de alta qualidade nesta língua. Finalmente, lidar com comentários em português, que frequentemente contêm erros gramaticais e ortográficos, adicionou uma camada extra de dificuldade ao processamento e análise

dos dados.

Os especialistas consultados avaliaram positivamente o método, destacando sua fácil compreensão e aplicação. Eles ressaltaram que a estrutura do método permite relacionar avaliações de produtos e serviços com os perfis dos usuários, o que pode ser utilizado para fornecer recomendações personalizadas. Também foi enfatizada a importância da extração de tópicos, sentimentos e entidades nomeadas para uma melhor compreensão das percepções dos consumidores. Esses resultados são valiosos para melhorar a experiência do cliente (CX) e a qualidade dos produtos e serviços.

Para trabalhos futuros, algumas melhorias foram sugeridas. A utilização de dados "Ground Truth" robustos pode aumentar a confiabilidade das análises e do treinamento do modelo de análise de sentimento. Também é recomendada uma avaliação mais aprofundada de métricas para comprovar a eficácia do método, especialmente se for implementado em ambientes de produção. Além disso, explorar o uso de modelos de linguagem de grande porte (LLMs) pode melhorar a assertividade e escalabilidade do método.

Ademais, apresentação dos comentários relacionados às entidades, juntamente com os sentimentos associados a elas, poderia proporcionar insights valiosos sobre a percepção dos consumidores em relação a aspectos específicos dos produtos e serviços analisados. Nesse sentido, outra possível melhoria seria aprimorar a filtragem das entidades identificadas, categorizando-as de forma mais detalhada, como organizações, nomes próprios de pessoas, locais, entre outros. Isso poderia proporcionar uma análise mais refinada e específica, destacando com maior precisão os aspectos mais relevantes nas avaliações.

Outrossim, uma definição de critérios de coerência, consistência e informatividade entre o ChatGPT e o Gemini é importante, de forma que a escolha não seja definida por uma avaliação pessoal. Outra questão seria a análise da aplicação das técnicas do método em outras ordens, como por exemplo, ao invés de fazer a modelagem de tópicos, a análise de sentimentos e extração de entidades nomeadas por tópico, realizar a análise de sentimentos, a modelagem de tópicos e a extração de entidades nomeadas por sentimento.

Nesse contexto de análise, outro ponto seria a especialização da metodologia, aplicando-a de maneira distinta dependendo se os dados referem-se à compra de

produtos ou à prestação de serviços. Isso se deve às diferentes etapas pelas quais os consumidores passam em cada situação. Além disso, é importante realizar uma filtragem por dimensões específicas do objeto analisado. Por exemplo, ao analisar a compra de produtos, a metodologia poderia focar em avaliações relacionadas à entrega ou aos custos. Já no contexto de prestação de serviços, a aplicação poderia se concentrar em aspectos como o atendimento ou à qualidade.

Finalmente, ajustar o método para lidar com comentários multi-tópicos e aplicar o método regularmente, de forma mensal ou semanal, para incorporar novos tópicos emergentes que possam surgir com novas avaliações são passos importantes para garantir a relevância contínua do método.

14. REFERÊNCIAS BIBLIOGRÁFICAS

- AKULICK, Samantha et al. Intent detection through text mining and analysis. 2017.
- ANDREWS, Joe. What is Customer Sentiment Analysis and Ways It Improves the Customer Experience. **SupportLogic**, 2022. Disponível em: <<https://www.supportlogic.com/resources/blog/what-is-customer-sentiment-analysis-and-7-ways-it-improves-the-customer-experience/>>. Acesso em: 3 out. 2023.
- AVINASH, M.; SIVASANKAR, E. A study of feature extraction techniques for sentiment analysis. In: **Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS 2018, Volume 3**. Springer Singapore, 2019.
- B2W Reviews. **Americanas Tech**, 2021. Disponível em: <<https://github.com/americanas-tech/b2w-reviews01>>. Acesso em: 12 mar. 2024.
- BAFNA, Prafulla; PRAMOD, Dhanya; VAIDYA, Anagha. Document clustering: TF-IDF approach. In: **2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)**. IEEE, 2016.
- BARION, Eliana Cristina Nogueira; LAGO, Decio. Mineração de textos. **Revista de ciências exatas e tecnologia**, 2008.
- BIRKETT, Alex. How to Measure Customer Satisfaction in 8 Simple Steps. **HubSpot**, 2022. Disponível em: <<https://blog.hubspot.com/service/how-to-measure-customer-satisfaction>>. Acesso em: 2 out. 2023.
- BLEI, David M.; NG, Andrew Y.; JORDAN, Michael I. Latent dirichlet allocation. **Journal of machine Learning research**, 2003.
- CARO, Laura Martínez; GARCÍA, Jose Antonio Martínez. Measuring perceived service quality in urgent transport service. **Journal of Retailing and Consumer Services**, 2007.
- CRM: o que é e como estruturar. **BigDataCorp**, 2023. Disponível em: <<https://blog.bigdatacorp.com.br/2023/05/17/o-que-e-crm/>>. Acesso em: 2 out. 2023.
- DEVLIN, Jacob et al. Bert: Pre-training of deep bidirectional transformers for language understanding. **arXiv preprint arXiv:1810.04805**, 2018.
- DIAS, Fred. Brazilian Portuguese Sentiment Analysis Datasets. **Kaggle**, 2021.

Disponível em: <<https://www.kaggle.com/datasets/fredericods/ptbr-sentiment-analysis-datasets>>. Acesso em: 12 mar. 2024.

E-COMMERCE brasileiro lidera o ranking mundial de crescimento. **Mundo do Marketing**, 2023. Disponível em: <<https://www.mundodomarketing.com.br/e-commerce-brasileiro-lidera-o-ranking-mundial-de-crescimento/>>. Acesso em: 3 out. 2023.

FENG, Jun et al. Product feature extraction via topic model and synonym recognition approach. In: **CCF Conference on Big Data**. Singapore: Springer Singapore, 2019.

GROOTENDORST, Maarten. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. **arXiv preprint arXiv:2203.05794**, 2022.

HONG, Wei et al. Analyzing the relationship between consumer satisfaction and fresh e-commerce logistics service using text mining techniques. **Sustainability**, 2019.

HOTH, Andreas; NÜRNBERGER, Andreas; PAAß, Gerhard. A brief survey of text mining. **Journal for Language Technology and Computational Linguistics**, 2005.

HUSSEIN, Doaa Mohey El-Din Mohamed. A survey on sentiment analysis challenges. **Journal of King Saud University-Engineering Sciences**, 2018.

JOVIĆ, Alan; BRKIĆ, Karla; BOGUNOVIĆ, Nikola. A review of feature selection methods with applications. In: **2015 38th international convention on information and communication technology, electronics and microelectronics (MIPRO)**. IEEE, 2015.

JÚNIOR, Severino Domingos da Silva; COSTA, Francisco José. Mensuração e escalas de verificação: uma análise comparativa das escalas de Likert e Phrase Completion. **PMKT—Revista Brasileira de Pesquisas de Marketing, Opinião e Mídia**, 2014.

KALRA, Vaishali; AGGARWAL, Rashmi. Importance of Text Data Preprocessing & Implementation in RapidMiner. **ICITKM**, 2017.

KHERWA, Pooja; BANSAL, Poonam. Topic modeling: a comprehensive review. **EAI Endorsed transactions on scalable information systems**, 2019.

KUMAR, V.; MINZ, S. Feature selection: A literature review. **SmartCR**, 2014.

LYU, Fang; CHOI, Jaewon. The forecasting sales volume and satisfaction of organic

products through text mining on web customer reviews. **Sustainability**, 2020.

MACIEL, Thales Vaz et al. Mineração de dados em triagem de risco de saúde. **Revista Brasileira de Computação Aplicada**, 2015.

MIAO, Jianyu; NIU, Lingfeng. A survey on feature selection. **Procedia computer science**, 2016.

MODELO de linguagem. **Wikipédia: a enciclopédia livre**. Disponível em: <https://pt.wikipedia.org/wiki/Modelo_de_linguagem>. Acesso em: 9 jun. 2024.

MOHIT, Behrang. Named entity recognition. In: **Natural language processing of semitic languages**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014.

O que é RAG? **AWS Amazon**. Disponível em: <<https://aws.amazon.com/pt/what-is/retrieval-augmented-generation/>>. Acesso em: 9 jun. 2024.

OLIVEIRA, Alcione et al. Brazilian Portuguese Hotel's Reviews Corpus. **Mendeley Data**, 2021. Disponível em: <<https://data.mendeley.com/datasets/2w3kvrg97m/1>>. Acesso em: 15 mar. 2024.

Opinion Mining for Portuguese. **Opinando**, 2019. Disponível em: <<https://sites.google.com/icmc.usp.br/opinando/>>. Acesso em: 14 mar. 2024.

PATEL, Ravikumar; PASSI, Kalpdrum. Sentiment analysis on twitter data of world cup soccer tournament using machine learning. **IoT**, 2020.

QIZI SAMIYEVA, Maftuna Faxriddin et al. Text mining and it is development stages. **Science and Education**, 2023.

RAJANAK, Yutika; PATIL, Ratna; SINGH, Yadvendra Pratap. Language Detection Using Natural Language Processing. In: **2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)**. IEEE, 2023.

RUDOLF, Silas. RoBERTa Large NER. **Hugging Face**, 2023. Disponível em: <<https://huggingface.co/51la5/roberta-large-NER>>. Acesso em: 16 abr. 2024.

SALAU, Ayodeji Olalekan; JAIN, Shruti. Feature extraction: a survey of the types, techniques, applications. In: **2019 international conference on signal processing and communication (ICSC)**. IEEE, 2019.

SILVA, Fernando Medeiros. Proposta de um índice de avaliação da satisfação do uso do transporte público do município de São Carlos por meio da adaptação do Customer

Satisfaction Score. 2022.

SILVA, Maleesha De. (2023). Preprocessing Steps for Natural Language Processing (NLP): A Beginner's Guide. **Medium**. Disponível em: <<https://medium.com/@maleeshadesilva21/preprocessing-steps-for-natural-language-processing-nlp-a-beginners-guide-d6d9bf7689c9>>. Acesso em: 18 nov. 2023.

SOUZA, Fábio; NOGUEIRA, Rodrigo; LOTUFO, Roberto. BERTimbau: pretrained BERT models for Brazilian Portuguese. In: **Intelligent Systems: 9th Brazilian Conference, BRACIS 2020, Rio Grande, Brazil, October 20–23, 2020, Proceedings, Part I 9**. Springer International Publishing, 2020.

TURCHI, Sandra R. Estratégias de marketing digital e e-commerce. **Editora Atlas SA**, 2000.

WEBER, Simone. Satisfação e fidelização dos docentes: um estudo no Centro Sinodal de Ensino Médio Dorothea Schäfer. 2021.

WEI, Chih-Ping et al. Understanding what concerns consumers: a semantic approach to product feature extraction from consumer reviews. **Information Systems and E-Business Management**, 2010.

WHAT is text mining? **IBM**, 2023. Disponível em: <<https://www.ibm.com/topics/text-mining>>. Acesso em: 6 nov. 2023.

WHAT is text mining? **MonkeyLearn**, 2023. Disponível em: <<https://monkeylearn.com/text-mining/>>. Acesso em: 6 nov. 2023.

YANG, YunYun et al. Text mining and visualization tools—impressions of emerging capabilities. **World Patent Information**, 2008.

YUAN, Lik Xun. Distilbert Base Multilingual Cased Sentiments Student. **Hugging Face**, 2023. Disponível em: <<https://huggingface.co/lxyuan/distilbert-base-multilingual-cased-sentiments-student>>. Acesso em: 14 mar. 2024.

ZHAO, Jun; LIU, Feifan. Product named entity recognition in Chinese text. **Language Resources and Evaluation**, 2008.

15. ANEXOS

Análise de Satisfação de Serviços e Produtos Brasileiros em Plataformas Eletrônicas: Uma Abordagem de Mineração de Texto

Isabella Lopes Carmo dos Santos
Ciência da Computação

Marcos Vianna Villas

Visão Geral

- Introdução
- Declaração do Problema
- Etapas do Método
- Aplicação em Diversos Domínios
- Caso de Estudo



Introdução

- Construção de um método para análise e extração de insights a partir das avaliações de produtos e serviços.
- Utilização de ferramentas de mineração de texto e inteligência artificial.

Declaração do Problema

Para o que exatamente o método serve?

Objetivos

- Mapeamento de assuntos e interesses dos clientes.
- Percepção dos sentimentos e satisfação dos clientes.
- Identificação de entidades relevantes.

Etapas do Método

1

Modelagem de Tópicos

- Identificação dos principais temas nas avaliações.
- Palavras-chave representam cada um dos tópicos.

2

Análise de Sentimentos

- Categorização das opiniões como positivas, neutras ou negativas.
- Nível de satisfação dos clientes por tópico.

3

Extração de Entidades Nomeadas

- Identificação de entidades relevantes (organizações, locais, pessoas).
- Associações com a experiência do cliente.

Aplicação em Diversos Domínios



Restaurantes

Como o Guacamole.



Hotéis

Como o Grand Hyatt.



Plataformas de Comércio Eletrônico

Como o Buscapé.

Caso de Estudo – Smartphone da Motorola

Avaliações coletadas no ano de 2013 da plataforma de comércio eletrônico da Buscapé contendo aproximadamente 84 mil avaliações.

Smartphone Motorola

O método foi aplicado ao produto "Smartphone Motorola Moto G 5S Dual Chip Android 7.1.1 Nougat Tela 5.2" Snapdragon 430 32GB 4G Câmera 16MP - Platinum", identificado no conjunto de dados da plataforma de comércio eletrônico da B2W por meio do nome do produto, através de uma coluna existente no conjunto, que é o mais avaliado nesse contexto. Para isso, os dados foram filtrados para recuperar apenas os comentários relacionados a esse produto em específico, totalizando 792 avaliações.

	Hotéis
Origem	Smartphone Motorola da plataforma de comércio eletrônico B2W
Tamanho	792 avaliações utilizadas
Período	Avaliações de Janeiro a Maio de 2018
Formato	Arquivo CSV
Colunas principais	Comentário da avaliação, estrelas da

Sessão de Perguntas

Perguntas

- Como você avalia a facilidade de compreensão do método
- Como você avalia o uso do método?
- Como você avalia a utilidade (relevância, impacto) em saber os principais assuntos discutidos nas avaliações dos clientes de uma empresa?
- Como você avalia a utilidade dos sentimentos extraídos pelo método?
- Como você avalia a utilidade das entidades nomeadas extraídas pelo método?
- Em sua opinião, o método fornece insights úteis para melhorar a qualidade de produtos e serviços?
- Na sua avaliação, o método pode ser aprimorado de alguma forma? Se sim, de que forma?