



**INSTITUTO
FEDERAL**

Amazonas

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO

AMAZONAS - CAMPUS DISTRITO

ENGENHARIA DE CONTROLE E AUTOMAÇÃO

PAULO HENRIQUE BARROS DINIZ

ANÁLISE DE SENTIMENTOS NO COMÉRCIO ONLINE UTILIZANDO BERTIMBAU

PARA O PORTUGUÊS BRASILEIRO

MANAUS - AM

2024

PAULO HENRIQUE BARROS DINIZ

ANÁLISE DE SENTIMENTOS NO COMÉRCIO ONLINE UTILIZANDO BERTIMBAU
PARA O PORTUGUÊS BRASILEIRO

Trabalho de Conclusão de Curso (TCC) apresentado ao curso de Engenharia de Controle e Automação do INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO AMAZONAS - CAMPUS DISTRITO, como requisito parcial para obtenção do grau de bacharel em Engenharia de Controle e Automação.

Orientador: Prof. Me. Marlos André
Silva Rodrigues.

MANAUS - AM

2024

Dados Internacionais de Catalogação na Publicação (CIP)

D585a Diniz, Paulo Henrique Barros.
Análise de sentimentos no comércio online utilizando BERTimbau para o português brasileiro / Paulo Henrique Barros Diniz. — Manaus, 2024. 52f.: il. color.

Monografia (Graduação) — Instituto Federal de Educação, Ciência e Tecnologia do Amazonas, *Campus* Manaus Distrito Industrial, Curso de Engenharia de Controle e Automação, 2024.

Orientador: Prof.º Marlos André Silva Rodrigues, Me.

1. BERTimbau. 2. Processamento de linguagem natural. 3. Comércio eletrônico. 4. Produtos. I. Rodrigues, Marlos André Silva. II. Instituto Federal de Educação, Ciência e Tecnologia do Amazonas. III. Título.

CDD 629.89

ANEXO 7

ATA DE DEFESA PÚBLICA DO TRABALHO DE CONCLUSÃO DE CURSO

Aos 19 dias do mês de março, de 2024, às 20:40 h, o(a) discente Paulo Henrique Barros Diniz apresentou o seu Trabalho de Conclusão de Curso para avaliação da Banca Examinadora constituída pelos seguintes integrantes: Prof(a). Marlos André Silva Rodrigues (docente-orientador), Prof(a). Alyson de Jesus dos Santos (coorientador), Prof(a). Gilbert Breyes Martins (Membro 1) e Prof(a). Gilbert Breyes Martins (Membro 2). A sessão pública de defesa foi aberta pelo(a) presidente da banca, que apresentou a Banca Examinadora e deu continuidade aos trabalhos, fazendo uma breve referência ao TCC, que tem como título ANÁLISE DE SENTIMENTOS NO COMÉRCIO ONLINE UTILIZANDO O ALGORITMO DE OPORTUNIDADES. Na sequência, o(a) discente teve até 30 minutos para a comunicação oral de seu trabalho. Cada integrante da banca examinadora fez suas arguições após a defesa do mesmo. Ouvidas as explicações do(a) discente, a banca examinadora, reunida em caráter sigiloso, para proceder à avaliação final, deliberou e decidiu pela APROVAÇÃO com média final 9,0 (NOVE)

do referido trabalho.

Foi dada ciência ao(à) discente que a versão final do trabalho deverá ser entregue até o dia 30 / 04 / 2024, com as devidas alterações sugeridas pela banca. Nada mais havendo a tratar, a sessão foi encerrada às 22 h 00 min, sendo lavrada a presente ata, que, uma vez aprovada, foi assinada por todos os membros da Banca Examinadora e pelo(a) discente.

Prof.(a) Orientador(a)/Presidente: Paulo Henrique Barros Diniz

Prof.(a) Avaliador 1: Alyson de Jesus dos Santos

Prof.(a) Avaliador 2: Gilbert Breyes Martins

Discente: PAULO HENRIQUE BARROS DINIZ

AGRADECIMENTOS

Primeiramente, gostaria de expressar minha mais profunda gratidão a Deus, por me guiar com sabedoria e força ao longo desta jornada. Sua presença constante foi a luz que iluminou os caminhos mais desafiadores, permitindo-me superar cada obstáculo com fé e perseverança.

À minha mãe, o pilar da minha vida, minha eterna gratidão. Sua inabalável crença em mim, seu amor e apoio incondicionais foram os alicerces que me sustentaram nos momentos mais difíceis. Mãe, você é a verdadeira heroína desta conquista, e tudo que eu sou e serei é graças a você.

À minha namorada, Jayne, um enorme obrigado por estar ao meu lado com tanto entusiasmo e apoio. Seu apoio me deu a confiança necessária para seguir em frente. Obrigado por ser essa força positiva e por tornar tudo mais divertido.

Aos meus amigos Ytalo Ribeiro e Rafael Reinaldo, por toda a companheirismo, suporte e momentos de descontração que me permitiram manter o equilíbrio e a sanidade em meio às demandas acadêmicas.

Aos meus amigos Guilherme, Vitor, Clara, Rubens e Maria, cada um de vocês contribuiu de forma única para essa conquista, seja através de palavras de incentivo ou ajudando nos estudos.

À minha família extensa, especialmente meus tios, tias e primas, por seu apoio incondicional. Vocês sempre acreditaram em mim e me encorajaram a seguir meus sonhos, proporcionando não apenas suporte emocional, mas também aqueles momentos de alegria e descontração que tanto são necessários.

Gostaria de expressar minha gratidão especial ao Prof. Dr. Alyson de Jesus dos Santos e ao Me. Marlos Andre Silva Rodrigues, por sua orientação excepcional, sabedoria e paciência. Vocês não apenas me guiaram academicamente, mas também me inspiraram a buscar excelência em minha jornada. Seu apoio foi fundamental para o meu desenvolvimento e sucesso.

Ao Instituto Federal do Amazonas, minha gratidão pela excelente formação, recursos e ambiente que propiciaram meu crescimento intelectual e pessoal. Esta instituição não apenas me forneceu o conhecimento técnico necessário, mas também me ensinou valores importantes que levarei adiante em minha carreira profissional.

A todos vocês, meu mais sincero agradecimento. Esta conquista não é apenas minha, mas nossa, compartilhada com cada um que esteve ao meu lado nesta jornada. Obrigado por fazerem parte da minha história.

"A educação é a arma mais poderosa que você
pode usar para mudar o mundo."

(Nelson Mandela)

RESUMO

A era digital, com o aumento do comércio eletrônico, transformou a forma como os consumidores e as empresas interagem. Embora as compras online ofereçam conveniência e rapidez, também criaram uma barreira física entre consumidores e empresas. Isso levou a uma necessidade crescente de métodos de comunicação digital eficazes. As avaliações online tornaram-se uma ferramenta essencial neste processo, funcionando como uma ponte de feedback entre consumidores e empresas. As opiniões compartilhadas online podem influenciar diretamente as decisões de compra de outros consumidores. Além disso, esses comentários fornecem insights para as empresas melhorarem seus produtos e serviços. Para mitigar esse distanciamento, as plataformas de comércio eletrônico incentivam o compartilhamento de opiniões e sentimentos dos usuários sobre os produtos, o que impacta nas decisões de compra de outros consumidores e fornece feedback para as empresas. Dada a grande quantidade e variedade de comentários, surge o desafio de analisá-los de forma eficiente. Aqui, o processamento de linguagem natural (PNL) desempenha um papel crucial. Ao integrar técnicas de PNL, como o modelo BERTimbau, é possível realizar análises automatizadas dos sentimentos expressos nas avaliações online. Os resultados experimentais mostram a eficácia do modelo, com precisões variando entre 75% a 88%, dependendo da categorização dos sentimentos e do tratamento das avaliações.

Keywords: BERTimbau; Processamento de linguagem natural; Sentimentos; Opinião; Produtos.

ABSTRACT

The digital age, with its rising e-commerce, has transformed the way consumers and businesses interact. While online shopping offers convenience and speed, it has also created a physical barrier between consumers and businesses. This has led to a growing need for effective digital communication methods. Online reviews have become an essential tool in this process, acting as a feedback bridge between consumers and businesses. The opinions shared online carry considerable weight, directly influencing the purchase decisions of other consumers. Moreover, these comments provide valuable insights for businesses to improve their products and services. To mitigate this distancing, e-commerce platforms encourage the sharing of opinions and feelings of users about products, which significantly impacts the purchasing decisions of other consumers and provides valuable feedback for businesses. Faced with the vast quantity and variety of comments, the challenge of analyzing them efficiently arises. The BERTimbau model emerges as a solution, offering automated sentiment analysis in Brazilian Portuguese. Experimental results demonstrate the effectiveness of the model, with accuracies ranging between 75% to 88% depending on the categorization of sentiments and treatment of 3-star reviews.

Keywords: BERTimbau; Natural Language Processing; Sentiments; Feedback; Products.

LISTA DE FIGURAS

Figura 1 – Uso da internet no Brasil de 2019 a 2023.	12
Figura 2 – Gráfico Lealdade dos Clientes.	13
Figura 3 – Brasileiros que Já Compraram Online.	17
Figura 4 – Ramos IA.	21
Figura 5 – Etapas Criação Modelo de Inteligência Artificial.	22
Figura 6 – Dispositivos Que Utilizam Assistentes Virtuais.	25
Figura 7 – Exemplo de tokenização.	26
Figura 8 – Arquitetura BERT.	29
Figura 9 – BERTimbau.	30
Figura 10 – Matriz de Confusão.	31
Figura 11 – Hugging Face.	33
Figura 12 – Kaggle.	34
Figura 13 – Resultado da Análise Cumulativa.	36
Figura 14 – Resultado da Análise dos Modelos de Classificação.	37
Figura 15 – Fluxo Metodologia.	39
Figura 16 – Modelo de Três Classes.	42
Figura 17 – Modelo Positivo/Negativo Simplificado.	43
Figura 18 – Modelo Negativo Ampliado.	43
Figura 19 – Exclusão de Neutros.	44
Figura 20 – Primeira Configuração com Valores Negativos, Neutros e Positivos.	46
Figura 21 – Segunda Configuração com Valores 1, 2 e 3 Negativos e 4 e 5 Positivos.	47
Figura 22 – Terceira Configuração com Valores 1 e 2 Negativos e 3, 4 e 5 Positivos.	48
Figura 23 – Quarta Configuração com Valores 1 e 2 Negativos, 4 e 5 Positivos e Sem Valores 3.	49
Figura 24 – Gráfico dos Resultados de Precisão.	50
Figura 25 – Gráfico dos Resultados de Recall.	50
Figura 26 – Gráfico dos Resultados de F1-score.	51

LISTA DE TABELAS

Tabela 1 – Descrição das siglas relacionadas ao treinamento do BERT.	28
Tabela 2 – Projetos Relacionados a Análise de Sentimentos.	37
Tabela 3 – Avaliações de Clientes no Buscapé.	40
Tabela 4 – Distribuição das Classificações das Avaliações.	41
Tabela 5 – Especificações Técnicas do Hardware Utilizado.	45
Tabela 6 – Hiperparâmetros Treinamento.	45
Tabela 7 – Relatório de Classificação da Primeira Configuração.	46
Tabela 8 – Relatório de Classificação Segunda Configuração.	47
Tabela 9 – Relatório de Classificação da Terceira Configuração.	48
Tabela 10 – Relatório de Classificação da Quarta Configuração	49
Tabela 11 – Projetos Relacionados a Análise de Sentimentos.	52

LISTA DE ABREVIATURAS E SIGLAS

AM	Aprendizado de Máquina
ARPANET	Advanced Research Projects Agency Network
BERT	Bidirectional Encoder Representations from Transformers
COVID-19	Coronavirus Disease 2019
FN	False Negative
FP	False Positive
GPT	Generative Pre-trained Transformer
IA	Inteligência Artificial
LN	Linguagem Natural
LR	Logistic Regression
MDIC	Ministério do Desenvolvimento, Indústria, Comércio e Serviços
NB	Naïve Bayes
PLN	Processamento de Linguagem Natural
RNNs	Redes Neurais Recorrentes
SVM	Support Vector Machine
TN	True Negative
TP	True Positive

SUMÁRIO

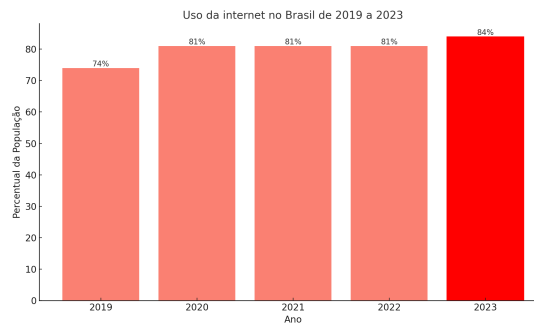
1	INTRODUÇÃO	12
1.1	Motivação	13
1.2	Problemática	14
1.3	Objetivos	15
1.3.1	<i>Objetivo Geral</i>	15
1.3.2	<i>Objetivos Específicos</i>	15
1.4	Justificativa	15
1.5	Organização do Texto	16
2	REFERENCIAL TEÓRICO	17
2.1	Internet e e-Commerce	17
2.2	Feedback do Cliente	18
2.3	Importância da Análise de Sentimentos	19
2.3.1	<i>Influência nas Decisões de Compra</i>	19
2.3.2	<i>Melhoria de Produtos e Serviços através da Análise de Sentimentos</i>	19
2.4	Inteligência Artificial	20
2.5	Aprendizado de Máquina	23
2.6	Processamento de Linguagem Natural	23
2.6.1	<i>Aplicações do PLN</i>	24
2.7	Tokens	25
2.7.1	<i>Tokenização: O Processo de Criação de Tokens</i>	26
2.7.2	<i>Tokenização em Modelos Avançados de PLN</i>	26
2.8	BERT	27
2.8.1	<i>Funcionamento do BERT</i>	27
2.8.2	<i>Aplicações do BERT</i>	29
2.9	BERTimbau	29
2.9.1	<i>Especificidades do Modelo</i>	30
2.10	Métricas de Avaliação	31
2.11	Hugging Face	33
2.12	Kaggle	34
3	TRABALHOS CORRELATOS	35

3.1	Análise de Sentimentos e Seus Desafios	35
3.2	Análise de Sentimento em Avaliações de Produtos em Redes Sociais . . .	36
3.3	Avaliações de Produtos Online em Português Brasileiro	36
3.4	BERT para Extração de Sentimentos em Língua Portuguesa	37
3.5	Tabela comparativa dos recursos oferecidos pelos projetos	37
4	METODOLOGIA	39
4.1	Base de Dados	40
<i>4.1.1</i>	<i>Pré-processamento</i>	<i>41</i>
4.2	<i>Tokenização</i>	44
4.3	Treinamento	44
5	RESULTADOS	46
6	CONCLUSÕES E TRABALHOS FUTUROS	52
	REFERÊNCIAS	53

1 INTRODUÇÃO

A internet causou uma transformação notável na vida cotidiana dos brasileiros, tornando-se um componente fundamental para milhões de pessoas. Wanderley *et al.* (2022) mostra como essa mudança é evidente ao observarmos os dados apresentados pelo Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação, conforme ilustrado no gráfico da Figura 1, onde revelou que 84% dos brasileiros acima de 10 anos, aproximadamente 152 milhões de indivíduos, já estavam utilizando a internet. Essa mudança tecnológica não só mudou a forma como as pessoas se comunicam e acessam informações, mas também transformou os hábitos de consumo, com a digitalização das compras e vendas de produtos.

Figura 1 – Uso da internet no Brasil de 2019 a 2023.



Fonte: Wanderley *et al.* (2022).

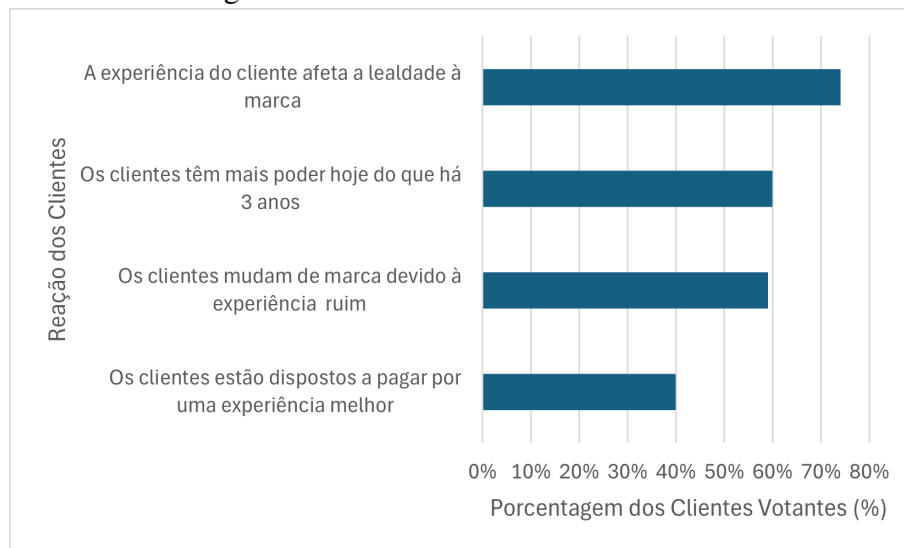
As vendas online no Brasil ganharam um impulso significativo durante a pandemia de Coronavirus Disease 2019 (COVID-19), resultando em uma ampla variedade de produtos e serviços oferecidos aos consumidores em diversas regiões do país. Segundo o Ministério do Desenvolvimento, Indústria, Comércio e Serviços (MDIC), entre 2020 e 2022, o comércio eletrônico brasileiro registrou um aumento expressivo, movimentando cerca de 450 bilhões de reais em operações de compra e venda. Essa tendência vem sendo explorada por grandes varejistas como a Amazon, que opera sem lojas físicas, e empresas nacionais que se expandiram fortemente no meio online, como Magazine Luiza e Americanas, com a venda de itens básicos até eletrônicos, vestuário e serviços, todos adquiridos com alguns cliques. Adicionalmente, Shimabukuro (2021) espera que o segmento de *e-commerce* no Brasil cresça em torno de 55% até 2024.

1.1 Motivação

A transição do consumo em lojas físicas para o comércio online alterou a dinâmica do *feedback* do cliente. Enquanto as lojas físicas proporcionavam uma interação direta e imediata, permitindo às empresas captar rapidamente as reações dos consumidores, o comércio eletrônico apresenta desafios diferentes. As plataformas de *e-commerce* permitem que os clientes adquiram produtos online, uma experiência que muitas vezes exclui aspectos sensoriais como tocar, cheirar ou provar.

Neste contexto digital, Sista *et al.* (2021) refere-se aos *feedbacks* dos clientes como uma ferramenta crucial de retorno para os fabricantes, uma vez que proporcionam um entendimento mais claro sobre as expectativas e necessidades dos consumidores. As opiniões expressas online são extremamente valiosas, por refletirem as experiências reais dos consumidores com produtos ou serviços. O gráfico da Figura 2 ilustra um estudo que demonstra o impacto da experiência do cliente na lealdade à marca até o ano de 2020, evidenciando que aproximadamente 74% da experiência influencia diretamente essa fidelidade (SUPEROFFICE, 2024).

Figura 2 – Gráfico Lealdade dos Clientes.



Fonte: SuperOffice (2024).

Essas informações fornecem percepções importantes que podem ser utilizados para melhorar a qualidade dos produtos, mudança de processos industriais, ajustes das estratégias de marketing e elevação da satisfação do cliente. Verhoef *et al.* (2009) contribui para essa compreensão, definindo a experiência do consumidor como um constructo multidimensional.

Este conceito engloba aspectos cognitivos, afetivos, emocionais, sociais e físicos da interação do cliente com o varejista. Essa definição ressalta a complexidade da experiência do comprador, especialmente no ambiente digital, onde a interpretação e a análise do *feedback* do cliente exigem uma compreensão abrangente dessas diversas dimensões.

1.2 Problemática

No cenário atual de competitividade acirrada no mercado, as avaliações de clientes, tornaram-se um elemento crucial para o sucesso e aprimoramento de produtos e serviços oferecidos pelas empresas e indústria. Com a ascensão das plataformas digitais e do *e-commerce*, consumidores estão cada vez mais propensos a expressar suas opiniões e experiências online, impactando diretamente a reputação das marcas, de acordo com Sista *et al.* (2021), elas desempenham um papel importante na identificação de sua satisfação com o produto. Neste contexto, o monitoramento e a análise dessas informações se tornam essenciais para as empresas compreenderem as expectativas e preferências de seus clientes, bem como identificar pontos de melhoria em seus produtos ou serviços, os quais impactam a produção diretamente na indústria.

Contudo, diante do volume massivo de dados gerados pelos consumidores, a análise manual de avaliações tornou-se um desafio monumental e impraticável. Neste cenário, a implementação de tecnologias de Processamento de Linguagem Natural (PLN) e aprendizado de máquina para automatizar a análise de sentimentos e opiniões em *reviews* revela-se essencial. Como destacado por Zhang *et al.* (2019), a análise de sentimentos é um problema de classificação que se baseia em dados textuais, sendo abordado com sucesso através da adaptação de métodos convencionais de aprendizado de máquina. Essa abordagem automatizada oferece às empresas a capacidade de compreender de forma mais rápida e eficiente o *feedback* dos consumidores, facilitando decisões estratégicas voltadas para a melhoria contínua de produtos e serviços e permitindo uma resposta mais ágil às tendências e exigências do mercado.

Portanto, para manter sua relevância e competitividade, as empresas precisam investir em tecnologias que lhes permitam filtrar, analisar e interpretar eficientemente a abundância de *feedbacks* dos consumidores disponíveis de forma online. Este investimento não apenas aprimora a percepção da marca perante o público, como também direciona inovações e melhorias alinhadas com as reais necessidades e desejos dos clientes.

1.3 Objetivos

1.3.1 *Objetivo Geral*

Investigar a aplicabilidade e eficácia do modelo BERTimbau e de técnicas avançadas de processamento de linguagem natural na análise de sentimentos expressos em avaliações online, visando extrair percepções que possam orientar estratégias empresariais e industriais para a melhoria de produtos.

1.3.2 *Objetivos Específicos*

- **Coletar e Preparar Dados:** extrair e preparar um conjunto de dados significativo de avaliações de clientes da plataforma Buscapé, garantindo a diversidade e representatividade das opiniões.
- **Explorar a Biblioteca *Hugging Face* para Tokenização:** utilizar as funcionalidades de tokenização e processamento de linguagem natural da biblioteca disponíveis na *Hugging Face* para estruturar e preparar os dados de *reviews* para análise.
- **Treinar Modelos com BERTimbau:** treinar modelos como a base de dados coletada e processada.
- **Avaliar o Desempenho e Precisão do Modelo:** verificar a precisão e eficácia do modelo BERTimbau e do processo de análise de sentimentos através de métricas de avaliação de desempenho.
- **Comparar o Desempenho do BERTimbau em Diferentes Arranjos de Datasets:** avaliar como o modelo BERTimbau se comporta ao analisar diferentes arranjos de datasets. Isso inclui comparar sua eficiência e precisão ao processar conjuntos de dados com características variadas.

1.4 Justificativa

Na era digital, especialmente no *e-commerce*, a dinâmica entre empresas e consumidores mudou drasticamente, com avaliações online se tornando influentes nas decisões de compra e fornecendo *feedback* essencial para o aprimoramento de produtos e serviços. A análise de sentimentos dessas avaliações, enriquecida por modelos avançados de processamento de linguagem natural como o BERTimbau, que é especializado na língua portuguesa, permite às empresas

entenderem profundamente as necessidades e percepções dos consumidores. Esta abordagem vai além da mera categorização de *feedbacks* como positivos ou negativos, capturando as emoções e subjetividades dos clientes, oferecendo *insights* valiosos para melhorar a experiência do cliente e o desempenho comercial. Ao se concentrar em avaliações no contexto brasileiro, este estudo destaca a importância de adaptar tecnologias globais para contextos locais, contribuindo para a diversidade e inclusão no campo da inteligência artificial.

1.5 Organização do Texto

O restante desta tese está organizado da seguinte forma. No Capítulo 2, é feita a referência para o desenvolvimento dos cenários propostos. No Capítulo 3 são apresentados os projetos relacionados utilizados como referência para o desenvolvimento dos cenários de aplicação propostos. No Capítulo 4 é mostrada a metodologia dos cenários de aplicação vislumbradas neste trabalho. No Capítulo 5 são mostrados os resultados do trabalho. O Capítulo 6 apresenta a conclusão e as direções futuras para o trabalho.

2 REFERENCIAL TEÓRICO

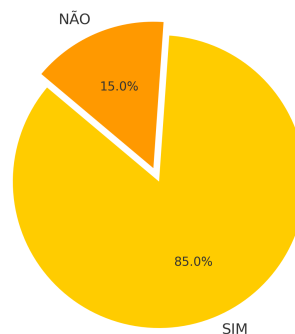
Neste capítulo, são apresentados aportes teóricos que deram subsídios à realização dessa pesquisa, referindo-se a internet, *e-commerce*, *feedback*, análise de sentimentos, inteligência artificial, aprendizado de máquina, processamento de linguagem natural, *tokens*, Bidirectional Encoder Representations from Transformers (BERT), BERTimbau, sobre o site da *Hugging Face* e *Kaggle*.

2.1 Internet e e-Commerce

A internet, conforme descrito por Leiner *et al.* (2009), é uma rede global de computadores interconectados que facilita a comunicação e troca de dados em todo o mundo. Originada na década de 1960 como Advanced Research Projects Agency Network (ARPANET), um projeto do Departamento de Defesa dos Estados Unidos da América, a internet evoluiu significativamente, refletindo não apenas avanços tecnológicos, mas também mudanças sociais impactantes, afetando desde a comunicação interpessoal até estruturas organizacionais globais.

No Brasil, segundo dados do Ministério das Comunicações (2023), 80% dos domicílios possuem acesso à internet, evidenciando a ampla democratização desse acesso. Esta expansão possibilitou o surgimento e crescimento do comércio eletrônico, uma forma de comércio que envolve a compra e venda de produtos e serviços pela internet. De acordo com Gonçalves (2024), 85% dos brasileiros com acesso à internet já realizaram compras online, demonstrando o enorme potencial do comércio eletrônico no país, conforme ilustrado no gráfico da Figura 3. Além disso, este crescimento é ainda mais potencializado pela integração com outras tecnologias, como a inteligência artificial abrindo caminho para novas oportunidades e inovações no setor.

Figura 3 – Brasileiros que Já Compraram Online.



Fonte: IAB Brasil e Offerwise (2023).

O *e-commerce*, ou comércio eletrônico, refere-se à compra e venda de bens ou serviços usando a internet, e à transferência de dinheiro e dados para executar essas transações. Com o advento da era digital, o *e-commerce* tornou-se um componente crucial da economia global. De acordo com a Statista (2020), as vendas globais de *e-commerce* devem atingir aproximadamente 4,8 trilhões de dólares americanos até 2021.

2.2 *Feedback* do Cliente

O *feedback* do cliente é reconhecido como uma fonte essencial de *insights* para empresas e pesquisadores de mercado, especialmente no contexto da inovação. De acordo com um estudo de Cabral e Marques (2023), a inovação em suas quatro dimensões (serviço, marketing, processo e organizacional) tem um impacto significativo na satisfação do cliente. Este estudo, focado na indústria hoteleira, destaca que a satisfação do cliente está associada à inovação, oferecendo *insights* valiosos para a tomada de decisão não apenas no setor hoteleiro, mas em diversos campos de atuação.

Além disso, com a evolução da tecnologia digital, a opinião do consumidor do cliente tornou-se mais interativo, multimodal e quase em tempo real, conforme discutido por Lee *et al.* (2023) em seu trabalho sobre aplicações de inteligência artificial em pesquisas de *feedback* do cliente. Este estudo, enfatiza que, embora a satisfação do cliente tenha sido anteriormente focado em pesquisas, ele evoluiu para incluir uma gama mais ampla de canais e formatos. A inteligência artificial e as tecnologias de aprendizado de máquina desempenham um papel crucial na compreensão e aproveitamento dessas fontes de dados de alta velocidade. No entanto, desafios como manipulação de avaliações e efeitos de rebanho permanecem, exigindo uma abordagem complementar de ferramentas de ciências sociais para extrair *insights* significativos desses dados.

No entanto, como destacado por Liu (2022), a extração de percepções úteis a partir dessas opiniões online representa um desafio significativo. O volume massivo de dados disponíveis em diversas fontes torna a tarefa de compilar e interpretar essas informações complexa para o observador médio. Por isso, a automação da análise de sentimentos, utilizando técnicas avançadas de processamento de linguagem natural e inteligência artificial, tornou-se uma necessidade crucial.

Além disso, a análise de *feedback* não se limita apenas à coleta de dados. Como apontado por Iacobucci *et al.* (2019), é essencial integrar e interpretar esses dados dentro do contexto mais amplo da jornada do cliente. Isso inclui compreender não apenas o que os clientes

estão dizendo, mas também por que estão dizendo, requerendo uma abordagem mais holística e empática.

Portanto, a verificação de *feedback* e principalmente dos sentimentos dos clientes é uma prática que combina tecnologia, psicologia do consumidor e estratégia de negócios, essencial para o sucesso e a adaptação das empresas no dinâmico mercado atual.

2.3 Importância da Análise de Sentimentos

A análise de sentimentos, é um subcampo do processamento de linguagem natural, a qual vital para o *e-commerce* devido ao seu impacto direto na percepção da marca e nas decisões de compra dos consumidores.

2.3.1 Influência nas Decisões de Compra

A influência das avaliações e comentários de clientes nas decisões de compra é um fenômeno crescente na era digital, redefinindo a maneira como consumidores interagem com produtos e serviços no mercado. Este cenário digital transformou as opiniões expressas online em elementos cruciais que afetam diretamente as escolhas de compra. Com a ascensão da análise de sentimentos, as empresas agora têm ferramentas mais sofisticadas para decifrar e utilizar estas opiniões a seu favor. Alguns pontos sobre decisões de compra são destacados a seguir:

- **Avaliações e Comentários de Clientes:** as opiniões expressas online influenciam fortemente as decisões de compra dos consumidores. Segundo um estudo, 93% dos consumidores afirmam que as avaliações online afetam suas decisões de compra (HANSEL, 2024).
- **Sentimentos como Indicadores:** a análise de sentimentos ajuda as empresas a entender melhor as opiniões dos clientes, segmentando-as em categorias como positiva, negativa ou neutra.

2.3.2 Melhoria de Produtos e Serviços através da Análise de Sentimentos

A análise de sentimentos extraída de avaliações e *feedbacks* online oferece às empresas *insights* valiosos sobre como os clientes percebem seus produtos e serviços. Essas informações são cruciais para diversas áreas estratégicas, incluindo:

- **Aprimoramento de Produtos e Serviços:** a análise detalhada das opiniões dos clientes

permite às empresas identificar pontos de força e fraqueza em seus produtos ou serviços. Isso pode levar a melhorias significativas, como alterações no design do produto, aprimoramento de funcionalidades, aperfeiçoamento da experiência do usuário, ou melhorias no atendimento ao cliente.

- **Desenvolvimento de Estratégias de Marketing:** compreender os sentimentos dos clientes ajuda a identificar características do produto ou serviço que são altamente valorizadas, que podem ser enfatizadas em campanhas de marketing. Simultaneamente, ajuda a reconhecer áreas que necessitam de melhorias, auxiliando na criação de estratégias de comunicação mais eficazes.
- **Personalização da Experiência do Cliente:** a análise aprofundada dos sentimentos dos clientes facilita a personalização das interações e ofertas da empresa. Isso não apenas melhora a satisfação do cliente, mas também contribui para a construção de uma base de clientes mais leal e engajada.
- **Gerenciamento de Reputação Online:** a análise de sentimentos é fundamental para o monitoramento e a gestão da reputação online de uma marca. Ela permite que as empresas identifiquem e respondam prontamente a *feedbacks* negativos, mitigando potenciais danos à sua imagem e fortalecendo sua presença no mercado digital.

Em resumo, a análise de sentimentos é uma ferramenta poderosa que, quando aplicada corretamente, pode fornecer às empresas informações cruciais para melhorar seus produtos, serviços e estratégias de marketing, ao mesmo tempo em que aprimora a experiência e satisfação do cliente.

2.4 Inteligência Artificial

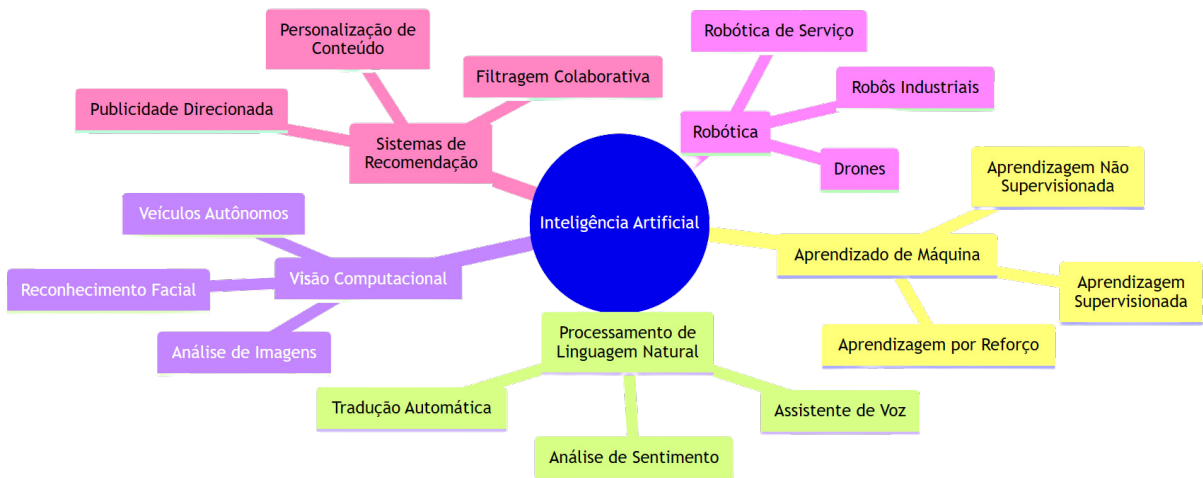
A Inteligência Artificial (IA) é um campo que tem sido amplamente definido na literatura, com cada definição enfatizando aspectos específicos de sistemas programados para exibir inteligência não-humana. Russell e Norvig (2010) descrevem IA como sistemas que replicam funções cognitivas tipicamente humanas, como aprendizado, fala e solução de problemas. Em contraste, Kaplan e Haenlein (2019) oferecem uma visão mais detalhada e elaborada, enfatizando a habilidade da IA de interpretar e aprender independentemente de dados externos. Essa aprendizagem permite adaptações flexíveis para alcançar objetivos específicos.

A evolução dessa área foi significativamente impulsionada pelo uso de *big data*, sendo o segmento do conhecimento que estuda como tratar, analisar e obter informações a

partir de conjuntos de dados muito grandes, facilitando o desempenho notável de algoritmos em tarefas específicas, como veículos robóticos, jogos e agendamento autônomo. Essa aplicação mais pragmática da IA contrasta com a abordagem mais cognitiva, de nível humano, onde complexidades como pensamento e emoções humanas ainda não foram completamente traduzidas em sistemas de IA (RUSSELL; NORVIG, 2010).

A era atual é marcada pela aplicação prática da inteligência artificial em diversos domínios. Esta tecnologia está revolucionando setores como a indústria automobilística, através do desenvolvimento de veículos autônomos, e o setor de saúde, com sistemas avançados de diagnóstico médico. No comércio eletrônico, ela aprimora a experiência do cliente com recomendações personalizadas e otimização de operações logísticas. Sua presença crescente influencia significativamente o modo de vida, trabalho e interação. A Figura 4 ilustra os diversos ramos que podem ser englobados pela IA.

Figura 4 – Ramos IA.



Fonte: Próprio Autor (2024)

Esta influência da inteligência artificial se estende ao setor de vendas, onde sua aplicação vai desde ferramentas de recomendação personalizadas até *chatbots* que interagem com os clientes em tempo real. Essas tecnologias estão transformando a jornada de compra, tornando-a mais eficiente e personalizada. Syam e Sharma (2018) exploram a influência de sistemas inteligentes nas vendas, com ênfase na melhoria da qualidade do serviço ao cliente e na fidelização. Ele detalha como, ao aprimorar a eficiência no atendimento ao cliente, a IA contribui significativamente para a satisfação e lealdade do consumidor, fornecendo *insights*

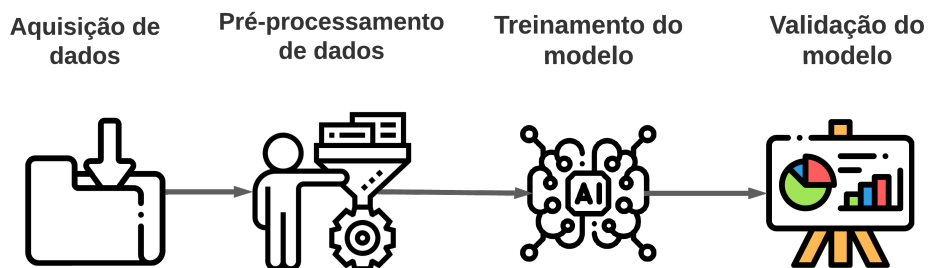
valiosos sobre a integração da IA no setor de vendas e seu impacto na experiência do cliente.

A sua aplicação impacta desde o início da jornada do cliente, superando a identificação e segmentação de potenciais clientes tradicionais por meio de algoritmos avançados que analisam padrões em dados de clientes e de mercado. Isso permite que as empresas analisem e identifiquem prospectos de alta qualidade, levando a campanhas de ativação de prospectos mais eficazes e personalizadas. Durante todo o processo de vendas, a IA oferece suporte contínuo e crítico, desde a proposta até o fechamento do negócio, analisando o comportamento do cliente, preferências e demografia para gerar conteúdo e mensagens personalizadas.

Após a aquisição de um novo cliente, pode ser aplicada para ajudar na integração e manutenção do mesmo. À medida que se avança na discussão sobre a inteligência artificial e suas aplicações práticas, é importante destacar o papel fundamental do aprendizado de máquina, um subcampo da área da IA, na evolução dessa tecnologia.

O procedimento para construir um modelo convencional de inteligência artificial é estruturado em quatro etapas fundamentais. Primeiro, ocorre a aquisição de dados, onde se coleta o conjunto de informações necessária para o treinamento do modelo. Segue o pré-processamento, fase que visa garantir a qualidade e adequação dos dados. A terceira etapa é o treinamento do modelo, onde os algoritmos aprendem a partir dos dados processados. Por fim, a validação do modelo é realizada para testar sua eficácia e precisão em condições reais ou simuladas. Esta última fase é realizada para garantir que o modelo funcione conforme esperado em aplicações práticas. Essas etapas estão ilustradas na Figura 5

Figura 5 – Etapas Criação Modelo de Inteligência Artificial.



Fonte: Próprio Autor (2024)

2.5 Aprendizado de Máquina

O Aprendizado de Máquina (AM) é caracterizado pelo uso de algoritmos avançados e análise de grandes volumes de dados para emular a capacidade humana de aprender com experiências. Segundo Jordan e Mitchell (2015), o AM tem se destacado por sua habilidade em classificar, prever e reconhecer padrões de maneira eficiente e inovadora. Este avanço tecnológico não apenas complementa as funcionalidades da IA em setores como vendas e atendimento ao cliente, mas também abre novas vertentes para aplicações ainda mais sofisticadas. Por meio da utilização desses algoritmos avançados, esta área busca replicar a capacidade humana de aprender a partir de experiências passadas e melhorar continuamente o desempenho em tarefas específicas.

Jordan e Mitchell (2015) destacam que a revolução do aprendizado de máquina permite que os computadores não apenas sigam instruções programadas, mas também adquiram conhecimento e habilidades de maneira autônoma e intuitiva. Isso amplia significativamente as fronteiras do que é possível em termos de automação e eficiência.

A capacidade de extrair visões significativas de grandes conjuntos de dados complexos é fundamental em muitos setores, desde a medicina até o marketing e a gestão de riscos financeiros. Ao empregar técnicas como redes neurais, aprendizado profundo e sistemas baseados em regras, o aprendizado de máquina está transformando a tomada de decisões, tornando-a mais informada, ágil e precisa. Com a crescente disponibilidade de dados e o avanço contínuo em algoritmos de aprendizado de máquina, espera-se que seu impacto continue crescendo, abrindo novos horizontes para inovações e aplicações em diversas áreas da sociedade.

2.6 Processamento de Linguagem Natural

A Linguagem Natural (LN) é a linguagem que utilizamos em nossa comunicação diária, seja na forma falada ou escrita, sendo caracterizada por sua complexidade, nuances e variações culturais. Diferentemente das linguagens de programação, que são precisas e estruturadas, a LN é inerentemente ambígua e flexível, variando significativamente entre diferentes idiomas, regiões e grupos culturais. A evolução constante é outra característica marcante, com novas palavras e expressões sendo continuamente adicionadas ao nosso léxico, enquanto outras caem em desuso. A dependência de contexto é fundamental, com o significado de palavras e frases mudando significativamente com base no contexto em que são empregadas (JURAFSKY;

MARTIN, 2009).

A compreensão da linguagem natural é crucial para diversas aplicações práticas no mundo da tecnologia, permitindo a criação de sistemas de inteligência artificial que podem interagir com humanos de maneira mais natural e intuitiva. Isso inclui assistentes virtuais, sistemas de tradução automática, ferramentas de análise de sentimentos e robôs de conversa automatizada. Essa evolução tecnológica é impulsionada pelo avanço contínuo no campo do processamento de linguagem natural.

O PLN é uma ramificação crucial da inteligência artificial e da Ciência da Computação, que visa capacitar os computadores para entender, interpretar e manipular a linguagem humana. O PLN combina disciplinas como linguística computacional, inteligência artificial e análise de dados para permitir que as máquinas leiam, compreendam e respondam à linguagem humana de uma maneira útil e significativa (AI, 2023).

No cerne do PLN estão os algoritmos e técnicas que permitem aos computadores processar e analisar grandes quantidades de dados de linguagem natural. Isso inclui a capacidade de reconhecer padrões de fala e escrita, compreender a semântica e a sintaxe, e extrair significado de contextos linguísticos complexos (MANNING; SCHUTZE, 1999). Os desafios do PLN abrangem desde a análise de sentimentos e a classificação de texto até a tradução automática e a geração de linguagem natural.

2.6.1 Aplicações do PLN

As aplicações do PLN são vastas e variadas, impactando diversos setores como *e-commerce*, saúde, mídia e educação. Alguns exemplos notáveis incluem assistentes virtuais, como Siri e Alexa (Figura 6), que utilizam PLN para entender e responder a comandos de voz, sistemas de tradução automática como o *Google Translate*, e ferramentas de análise de sentimentos que ajudam as empresas a compreender melhor as opiniões dos clientes sobre seus produtos e serviços (HAPKE *et al.*, 2019).

Figura 6 – Dispositivos Que Utilizam Assistentes Virtuais.



Fonte: Kinast (2023)

A ambiguidade da linguagem humana, as nuances culturais e as variações dialetais são aspectos que complicam a compreensão precisa da linguagem natural pelos computadores (BATES *et al.*, 1993). Além disso, a necessidade de grandes conjuntos de dados anotados para treinar modelos de PLN eficazes representa um desafio em termos de recursos e tempo (DENG; LIU, 2018).

Com o surgimento de modelos de aprendizado profundo e redes neurais, o campo do PLN tem testemunhado uma evolução significativa. Modelos como BERT (Bidirectional Encoder Representations from Transformers), GPT (Generative Pre-trained Transformer) e outros baseados em transformadores têm revolucionado a compreensão da linguagem natural, oferecendo avanços em tarefas como a compreensão de texto e a geração de linguagem (MEHTA, 2023).

Devido à alta complexidade e variação da linguagem natural, o pré-processamento de dados se torna uma etapa crucial no processamento de linguagem natural. Nesse estágio, ocorre a limpeza de dados, removendo elementos que não agregam valor significativo ao processamento, e a tokenização de textos. Sendo este último, o processo de dividir textos em unidades menores, ou *tokens*, facilitando a análise e interpretação por modelos de PLN.

2.7 Tokens

Um *token* pode ser entendido como uma instância de uma sequência de caracteres em um texto particular que são agrupados como um elemento semântico útil para processamento. Na prática, *tokens* são frequentemente palavras (Figura 7), mas podem incluir pontuações, números ou outros símbolos, dependendo da necessidade específica da tarefa de PLN.

Figura 7 – Exemplo de tokenização.

M E U N O M E É H U G O C A L I X T O	
" M E U "	
" N O M E "	CADA PALAVRA
" É "	É UM TOKEN
" H U G O "	
" C A L I X T O "	

Fonte: Calixto (2021)

2.7.1 Tokenização: O Processo de Criação de Tokens

O processo de transformar um texto em *tokens* é conhecido como tokenização. Essa etapa prepara os dados brutos (texto) para análise e processamento por modelos de PLN. A tokenização envolve a segmentação de texto em palavras, frases ou outros elementos significativos.

Existem diferentes abordagens de tokenização, cada uma adequada para diferentes aplicações:

- **Tokenização por palavras:** divide o texto em palavras individuais.
- **Tokenização por sentenças:** separa o texto em sentenças completas.
- **Tokenização subpalavra:** usada em modelos como o BERT, onde palavras são divididas em subcomponentes menores, permitindo que o modelo capture melhor a morfologia das palavras.

2.7.2 Tokenização em Modelos Avançados de PLN

O BERT, por exemplo, utiliza uma abordagem de tokenização subpalavra, que permite ao modelo lidar de forma mais eficaz com um vocabulário extenso, incluindo palavras desconhecidas ou raras. Os *tokens* são, portanto, a espinha dorsal do processamento de linguagem natural. Eles representam o ponto de partida para a maioria das tarefas de PLN e são fundamentais para a eficácia dos modelos modernos de PLN, como o BERT, em sua capacidade de interpretar e processar linguagem humana de forma significativa e precisa.

2.8 BERT

O BERT é um modelo de processamento de linguagem natural desenvolvido por Jacob Devlin e sua equipe na Google. Lançado em 2018, o BERT rapidamente estabeleceu-se como um marco na área, devido à sua capacidade de entender o contexto das palavras em um texto de maneira bidirecional (DEVLIN *et al.*, 2018).

2.8.1 Funcionamento do BERT

O BERT é baseado na arquitetura de transformadores, a qual foi introduzida pelo trabalho de Vaswani *et al.* (2017). O Transformer é um modelo baseado em atenção que dispensa a necessidade de Redes Neurais Recorrentes (RNNs) e convolucionais, focando inteiramente na mecanismo de atenção para processar os dados em paralelo, o que melhora significativamente a eficiência do treinamento.

Um dos principais avanços do BERT é sua capacidade bidirecional, o que significa que ele considera o contexto de ambos os lados (esquerdo e direito) de cada *token* (palavra ou subpalavra) no texto. Isso é uma diferença significativa em relação aos modelos anteriores, que geralmente analisavam o texto de forma unidirecional (da esquerda para a direita ou da direita para a esquerda). Essa abordagem bidirecional permite que o BERT tenha uma compreensão mais profunda do contexto e das nuances da linguagem, resultando em melhor desempenho em tarefas como compreensão de texto e inferência de linguagem.

O BERT foi pré-treinado simultaneamente em duas tarefas. A primeira delas, chamada de modelagem de linguagem, consiste em selecionar 15% dos *tokens* em um texto e prever o *token* escolhido com base em seu contexto (DEVLIN *et al.*, 2018). Os significados de cada sigla utilizada no treinamento do modelo estão presentes na Tabela 1. Visto isso, o *token* escolhido poderia ser:

- Substituído por um *token* [MASK] com probabilidade de 80%.
- Substituído por um *token* de palavra aleatória com probabilidade de 10%.
- Não ser substituído com probabilidade de 10%.

Por exemplo, considere a frase "A bicicleta azul é rápida". Se o quarto *token* ("rápida") fosse escolhido para previsão, o modelo veria a entrada como "A bicicleta [MASK] é rápida" com 80% de probabilidade, "A bicicleta verde é rápida" com 10% de probabilidade e "A bicicleta azul é rápida" com 10% de probabilidade. Após processar a entrada, o quarto vetor de

saída do modelo era alimentado em uma rede neural separada, que gerava uma distribuição de probabilidade em relação ao vocabulário de 30.000 palavras.

Já a tarefa 2 é intitulada de precisão da próxima frase, na qual o modelo tinha que prever se duas extensões de texto ocorriam sequencialmente no corpus de treinamento. Isso era feito atribuindo [IsNext] se as duas extensões estavam em sequência, ou [NotNext] se não estavam. O primeiro trecho começava com um *token* especial [CLS] (para "classificar"), e os dois trechos eram separados por um *token* especial [SEP] (para "separado"). Após processar os dois trechos, o primeiro vetor de saída (a codificação vetorial para [CLS]) era alimentado em uma rede neural separada para realizar uma classificação binária em [IsNext] e [NotNext].

Por exemplo, se tivéssemos a entrada "[CLS] O céu está claro [SEP] O sol está brilhando", o modelo deveria gerar o *token* [IsNext], indicando que as duas extensões ocorriam sequencialmente. No entanto, se a entrada fosse "[CLS] A música está alta [SEP] Qual é o clima hoje?", o modelo deveria gerar o *token* [NotNext], indicando que as extensões não ocorriam em sequência.

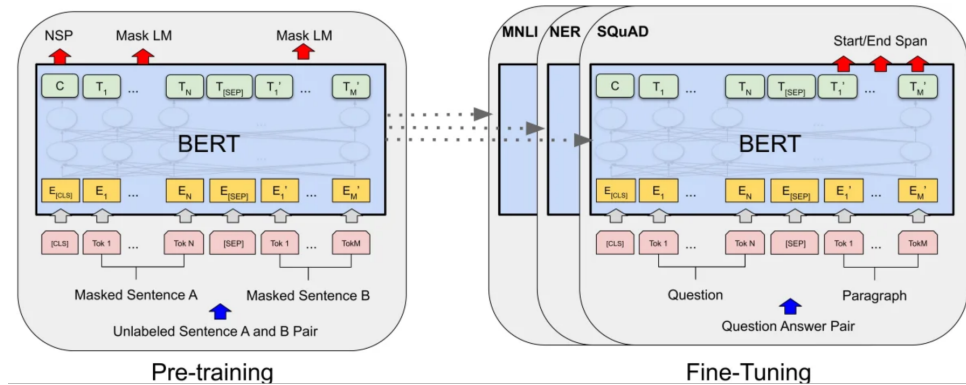
Sigla	Descrição
CLS	Token especial para "classificar"
SEP	Token especial para "separado"
MASK	Token usado para mascarar palavras durante o pré-treinamento
IsNext	Indica que duas extensões ocorrem sequencialmente
NotNext	Indica que duas extensões não ocorrem sequencialmente

Tabela 1 – Descrição das siglas relacionadas ao treinamento do BERT.

Essa abordagem bidirecional do BERT o tornou uma referência no campo de processamento de linguagem natural e análise de texto, ao melhorar significativamente o desempenho em uma variedade de tarefas, como análise de sentimentos, tradução automática e de resposta a perguntas.

Na etapa de ajuste fino (*fine-tuning*), o modelo pré-treinado é ajustado para tarefas específicas de PLN, como classificação de texto, análise de sentimentos, ou perguntas e respostas. Este processo envolve o treinamento do modelo em um conjunto de dados específico da tarefa, permitindo que ele se adapte a essa aplicação particular. A Figura 8 ilustra as fases de pré-treinamento e ajuste fino do modelo BERT, destacando os componentes-chave e processos envolvidos na adaptação do modelo para aplicações específicas de PLN.

Figura 8 – Arquitetura BERT.



Fonte: MindSpore (2022)

2.8.2 Aplicações do BERT

Desde o seu lançamento, o BERT transformou o campo do PLN, estabelecendo novos padrões para uma variedade de tarefas de compreensão da linguagem. Modelos subsequentes, como o RoBERTa, DistilBERT e o BERTimbau, foram inspirados pelo BERT, buscando aprimorar sua eficiência ou reduzir requisitos computacionais (SANH *et al.*, 2019). Tendo sido utilizado em uma ampla gama de tarefas de PLN, incluindo:

- **Compreensão de Leitura:** capacidade de responder perguntas sobre um dado texto.
- **Classificação de Texto:** determinar se uma avaliação de produto é positiva ou negativa.
- **Tradução Automática:** melhorar a precisão da tradução de texto entre idiomas.
- **Análise de Sentimentos:** avaliar a polaridade (positiva, negativa, neutra) das opiniões expressas em um texto.

Apesar de seu sucesso, o BERT possui limitações, incluindo a necessidade de grandes volumes de dados para o treinamento e desafios na interpretação de linguagens com estruturas gramaticais complexas ou pouco comuns. Além disso, a computação intensiva necessária para treinar e executar modelos baseados no BERT pode ser um obstáculo para sua aplicação em dispositivos com recursos limitados (DEVLIN *et al.*, 2018).

2.9 BERTimbau

O BERTimbau (Figura 9) é uma adaptação específica do modelo de linguagem BERT para o português brasileiro. Desenvolvido com o objetivo de preencher uma lacuna significativa no campo do processamento de linguagem natural para idiomas além do inglês, o BERTimbau

representa um avanço notável para pesquisadores e profissionais que trabalham com textos em português/br (SOUZA *et al.*, 2020).

Figura 9 – BERTimbau.



Fonte: Souza *et al.* (2020)

2.9.1 Especificidades do Modelo

A criação do BERTimbau envolveu o treinamento do modelo utilizando um corpus extenso e diversificado, composto por textos de diferentes gêneros e estilos, refletindo as nuances e a riqueza do português falado no Brasil. Esse processo de treinamento especializado permitiu que o modelo adquirisse uma compreensão mais aprofundada e matizada do idioma, comparativamente aos modelos multilíngues tradicionais (SOUZA *et al.*, 2020). O BERTimbau contribui para o campo do PLN de várias maneiras significativas como:

- **Compreensão Contextual Aprimorada:** ao ser treinado especificamente na língua portuguesa, o BERTimbau oferece uma compreensão contextual mais rica e precisa, essencial para interpretar corretamente o sentido de palavras e frases em diferentes contextos.
- **Flexibilidade e Versatilidade:** assim como o BERT original, o BERTimbau é flexível e pode ser adaptado para uma variedade de tarefas de PLN, incluindo mas não se limitando a análise de sentimentos, classificação de texto, e extração de informações.

Embora o BERTimbau represente um avanço significativo, ele não está isento de desafios. A necessidade de conjuntos de dados ainda mais amplos e a adaptação para contextos específicos são áreas que requerem atenção contínua. Além disso, a adaptação do modelo para outras variantes da língua portuguesa, como a falada em Portugal e em países africanos lusófonos, representa uma oportunidade de pesquisa futura interessante.

O BERTimbau se destaca como um recurso valioso para o avanço do PLN em

português brasileiro, oferecendo uma ferramenta robusta e eficaz para uma ampla gama de aplicações. Seu desenvolvimento é um passo importante para garantir que os avanços no campo do PLN sejam inclusivos e abrangentes, abordando as necessidades de comunidades linguísticas além do inglês.

2.10 Métricas de Avaliação

Métricas de avaliação são fundamentais no campo da inteligência artificial, pois permitem mensurar o desempenho e a eficácia dos modelos e algoritmos desenvolvidos. Elas variam de acordo com a aplicação específica da IA, mas algumas são amplamente utilizadas em diferentes domínios. A seguir é apresentado as métricas que serão utilizadas nesse trabalho.

- **Matriz de Confusão (*Confusion Matrix*):** a matriz de confusão é criada utilizando os resultados fornecidos por um algoritmo de classificação. Nela, tanto os acertos quanto os erros são contabilizados e acumulados, e esta contagem é organizada na forma de uma matriz, como representado na Figura 10.

Figura 10 – Matriz de Confusão.

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

Fonte: Nogare (2020)

- **Verdadeiro Positivo/True Positive (TP):** valores que o modelo identificou corretamente como positivos, correspondendo ao que era positivo na amostra (previsão acertada).
- **Verdadeiro Negativo:/True Negative (TN)**valores que o modelo classificou corretamente como negativos, alinhando-se com sua classificação negativa na amostra (previsão assertiva).
- **Falso Positivo/False Positive (FP):** valores que o modelo classificou erroneamente como positivos, enquanto na realidade eram negativos na amostra (falha na previsão).
- **Falso Negativo/False Negative (FN):** valores que foram incorretamente previstos

como negativos pelo modelo, mas que eram positivos na amostra (erro na previsão).

- **Acurácia (*Accuracy*):** é uma das métricas mais simples e comuns, especialmente em tarefas de classificação, representando a proporção de previsões corretas feitas pelo modelo em relação ao total de previsões. Apesar de ser fácil de entender e calcular, a acurácia pode ser enganosa em conjuntos de dados desequilibrados, onde uma classe é muito mais prevalente que as outras. Sua fórmula está ilustrado na Equação 2.1.

$$\text{Acurácia} = \frac{\text{Verdadeiros Positivos} + \text{Verdadeiros Negativos}}{\text{Total de Previsões}} \quad (2.1)$$

- **Precisão (*Precision*):** a precisão mede a exatidão das previsões positivas de um modelo de classificação. Calculando a proporção de verdadeiros positivos em relação ao total de casos classificados como positivos, incluindo os falsos positivos. Ela útil em situações onde os falsos positivos têm consequências graves ou indesejadas, como na identificação de doenças ou no filtro de e-mails, onde é mais crítico evitar classificar incorretamente os casos negativos como positivos. Sua fórmula está ilustrado na Equação 2.2.

$$\text{Precisão} = \frac{\text{Verdadeiros Positivos}}{\text{Verdadeiros Positivos} + \text{Falsos Positivos}} \quad (2.2)$$

- **Sensibilidade (*Recall*):** a *recall* mede a capacidade de um modelo de classificação em identificar todos os casos positivos relevantes. Ela é calculada pela proporção de verdadeiros positivos em relação à soma dos verdadeiros positivos e falsos negativos. Esta métrica é essencial em contextos onde é crítico detectar todos os casos positivos, mesmo à custa de aceitar alguns falsos positivos, como em diagnósticos médicos, onde perder um caso positivo pode ter consequências sérias. Sua fórmula está ilustrado na Equação 2.3.

$$\text{Recall} = \frac{\text{Verdadeiros Positivos}}{\text{Verdadeiros Positivos} + \text{Falsos Negativos}} \quad (2.3)$$

- **Pontuação F1 (*F1-score*):** a pontuação F1 é a média harmônica entre precisão e revocação, fornecendo um único valor para facilitar a comparação entre modelos. Esta métrica é útil quando é necessário balancear precisão e revocação, especialmente em casos de distribuição de classe desigual. Sua fórmula está ilustrado na Equação 2.4.

$$\text{F1-score} = 2 \times \frac{\text{Precisão} \times \text{Recall}}{\text{Precisão} + \text{Recall}} \quad (2.4)$$

2.11 *Hugging Face*

Hugging Face é uma organização que se tornou um nome proeminente no campo do processamento de linguagem natural, conhecida por sua contribuição significativa para o avanço das tecnologias de aprendizado de máquina e inteligência artificial. Sua plataforma oferece uma vasta gama de modelos de PLN pré-treinados e ferramentas que têm transformado a maneira como os pesquisadores e desenvolvedores abordam tarefas complexas de PLN.

Figura 11 – Hugging Face.



Fonte: Hugging Face (2024)

Ganhando notoriedade por sua biblioteca transformers, que fornece acesso fácil a modelos de PLN de última geração, como BERT, Generative Pre-trained Transformer (GPT), T5, entre outros. Esta biblioteca facilitou para os desenvolvedores a implementação de modelos complexos de PLN em suas aplicações, sem a necessidade de construí-los do zero. Algumas das características *Hugging Face* são:

- **Modelos Pré-treinados:** A plataforma oferece uma coleção extensiva de modelos pré-treinados, que podem ser facilmente customizados e ajustados para tarefas específicas de PLN.
- **Comunidade e Colaboração:** *Hugging Face* não é apenas uma biblioteca de ferramentas, mas também uma comunidade ativa onde pesquisadores, desenvolvedores e entusiastas de PLN colaboram e compartilham conhecimentos.
- **Facilidade de Uso:** Uma das maiores vantagens de é a sua facilidade de uso. Os desenvolvedores podem integrar modelos de PLN de alta complexidade com apenas algumas linhas de código.
- **Contribuição para a Pesquisa:** também contribui significativamente para a pesquisa em PLN, publicando regularmente estudos, papers e realizando experimentos que ajudam a avançar o campo.

Hugging Face tem desempenhado um papel crucial na democratização do acesso

a tecnologias de processamento avançadas. Isso permitiu que empresas de todos os tamanhos integrassem capacidades de PLN em seus produtos e serviços, melhorando a eficiência, a precisão e a capacidade de inovação em diversas aplicações, como chatbots, análise de sentimentos, sumarização automática e tradução automática.

2.12 *Kaggle*

Kaggle (Figura 12) é uma plataforma pertencente ao Google LLC, desenhada especificamente para a comunidade de ciência de dados e aprendizado de máquina. Funciona como um ponto de encontro virtual onde profissionais e entusiastas dessas áreas podem explorar e compartilhar uma vasta gama de conjuntos de dados, bem como desenvolver e testar modelos em um ambiente web especializado. A plataforma facilita a interação entre usuários, permitindo que colaborem em projetos, compartilhem conhecimento técnico e participem de desafios de modelagem.

Figura 12 – Kaggle.



Fonte: Cube (2020)

Além de ser um recurso para o desenvolvimento profissional individual, *Kaggle* também promove uma forma de colaboração coletiva. Os participantes podem se juntar a competições que têm como objetivo encontrar soluções inovadoras para problemas complexos de ciência de dados. Essas competições variam em escopo e dificuldade, abrangendo desde questões práticas do dia-a-dia até desafios de pesquisa de ponta.

O ecossistema *Kaggle* também inclui fóruns de discussão e um sistema de notebooks, permitindo que os usuários executem código em Python e R diretamente na plataforma, o que facilita a análise exploratória de dados e a construção de modelos sem a necessidade de configuração de ambiente local. Essas ferramentas e recursos fazem do Kaggle uma plataforma robusta para aprendizado contínuo, experimentação e competição na vanguarda da ciência de dados e aprendizado de máquina. A base de dados utilizada nesse trabalho foi extraída da plataforma.

3 TRABALHOS CORRELATOS

Este capítulo apresenta os principais projetos relacionados ao problema tratado neste trabalho. Na revisão da literatura sobre análise de sentimentos aplicada ao *e-commerce*, percebe-se uma notável prevalência de estudos que empregam o processamento de linguagem natural para explorar uma ampla gama de aplicações. Esses estudos vão desde o monitoramento de sentimentos em redes sociais até a análise de tendências de mercado, demonstrando a versatilidade e a eficácia do PLN. No entanto, emerge uma lacuna específica no que tange à análise detalhada de sentimentos relacionados a avaliações de produtos em sites de comércio eletrônico, especialmente em contextos onde se utiliza a língua portuguesa.

Embora haja uma riqueza de pesquisa em línguas como inglês e mandarim, o estudo de avaliações de *e-commerce* na língua portuguesa está relativamente subdesenvolvido. Esta lacuna representa uma oportunidade substancial para pesquisa, considerando a crescente popularidade do comércio eletrônico em países lusófonos. Entender as nuances culturais e linguísticas que influenciam as avaliações dos consumidores em português/br é crucial para as empresas que buscam se conectar efetivamente com esse segmento de mercado.

Além disso, muitos estudos existentes se concentram em abordagens genéricas para a análise de sentimentos, frequentemente sem levar em conta as características únicas do português/br, como sua complexidade sintática e semântica. A adaptação de ferramentas de PLN especificamente voltadas para a língua portuguesa, que possam captar e interpretar de maneira acurada os sentimentos expressos em avaliações online, é fundamental. Estas ferramentas forneceriam insights mais precisos e culturalmente relevantes, beneficiando empresas e marcas que atuam no setor de *e-commerce*.

3.1 Análise de Sentimentos e Seus Desafios

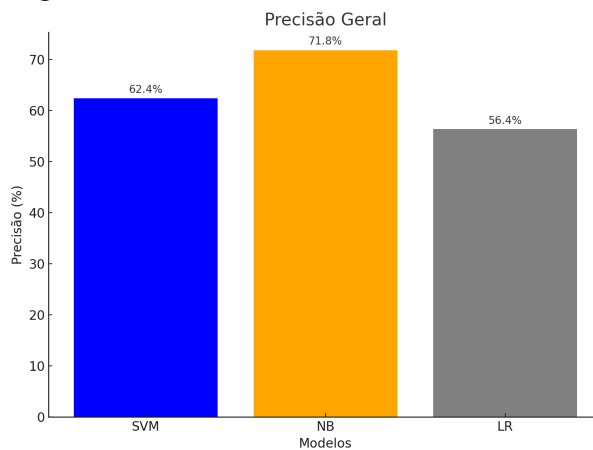
Pandey e Deorankar (2019), exploram os desafios da análise de sentimentos no PLN. Este estudo aborda o processo de análise de opiniões em textos, destacando as dificuldades em extrair entidades nomeadas, resolver co-referências e determinar a polaridade dos sentimentos. Ele também examina diferentes ferramentas usadas na análise de sentimentos, oferecendo uma matriz de avaliação para medir o desempenho dos sistemas de classificação de sentimentos.

3.2 Análise de Sentimento em Avaliações de Produtos em Redes Sociais

Dentro da área de análise de sentimentos online, o estudo conduzido por Nafees *et al.* (2018) apresenta contribui ao investigar a aplicação de técnicas de processamento de linguagem natural na classificação de avaliações de produtos em plataformas online. O estudo utilizou algoritmos como Naïve Bayes (NB), Support Vector Machine (SVM) e Logistic Regression (LR) para determinar a polaridade dos sentimentos expressos pelos usuários nas avaliações.

Os resultados obtidos apresentam que o SVM alcançou uma acurácia de 71,8%, enquanto o Naïve Bayes e a Regressão Logística obtiveram taxas de acerto de 56,4% e 62,4%, respectivamente. Os resultados estão expostos na tabela da Figura 13.

Figura 13 – Resultado da Análise Cumulativa.

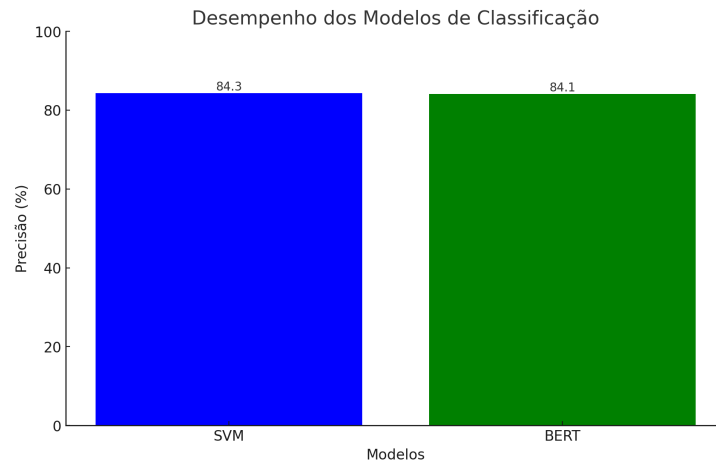


Fonte: Nafees *et al.* (2018)

3.3 Avaliações de Produtos Online em Português Brasileiro

Britto *et al.* (2023), foca em metodologias para identificar avaliações mais úteis para consumidores brasileiros, contribuindo para uma experiência de compra online mais informativa e satisfatória, a base de dados utilizada foi extraída do site de vendas Amazon, os resultados estão representados na tabela da Figura 14. O SVM obteve o melhor desempenho, alcançando 84,3% de precisão. Os modelos BERT tiveram um desempenho muito semelhante ao SVM, atingindo uma precisão de 84,1%. Apesar de terem alcançado um bom desempenho, esses classificadores apresentaram um longo tempo médio de execução para treinamento e teste, ambos levando mais de 13 minutos.

Figura 14 – Resultado da Análise dos Modelos de Classificação.



Fonte: Britto *et al.* (2023)

3.4 BERT para Extração de Sentimentos em Língua Portuguesa

Lopes *et al.* (2021) apresenta uma abordagem para a extração de aspectos usando modelos pré-treinados BERT, focando tanto em uma versão multilíngue (fornecida pelo Google) quanto em para a língua portuguesa (BERTimbau). Os experimentos relatados demonstram que a extração de aspectos baseada no BERT pré-treinado para o português-br alcançou uma acurácia de até 93% em um corpus de revisões sobre o setor de acomodações.

3.5 Tabela comparativa dos recursos oferecidos pelos projetos

A Tabela 11 apresenta o comparativo dos trabalhos de acordo com os seguintes aspectos: foco, tecnologia e a base de dados.

Tabela 2 – Projetos Relacionados a Análise de Sentimentos.

Trabalhos	Nafees <i>et al.</i> (2018)	Britto <i>et al.</i> (2023)	Lopes <i>et al.</i> (2021)	Trabalho Proposto
Foco	Redes Sociais	Site de Vendas	Site de Acomodações	Site de Vendas
Tecnologia	NB; SMV; LR	BERT; SVM	BERT; BERTimbau	BERTimbau
Base de Dados	Twitter	Amazon	Airbnb	Buscapé

Fonte: Próprio Autor (2024)

A comparação apresenta estudos no campo da análise de sentimentos e classificação de textos, abordando distintas fontes de dados e tecnologias de processamento de linguagem

natural. Os estudos de Nafees *et al.* (2018), Britto *et al.* (2023), Lopes *et al.* (2021), e o trabalho proposto exploram o uso de tecnologias como Naive Bayes , Support Vector Machine , Logistic Regression , e BERT (incluindo uma variante específica, o BERTimbau), aplicadas a bases de dados provenientes de plataformas como Twitter, Amazon, Airbnb, e Buscapé, respectivamente. Enquanto Nafees *et al.* (2018) foca na análise de sentimentos em redes sociais, Britto *et al.* (2023) e o trabalho proposto concentram-se em sites de vendas, e Lopes *et al.* (2021) em um site de acomodações.

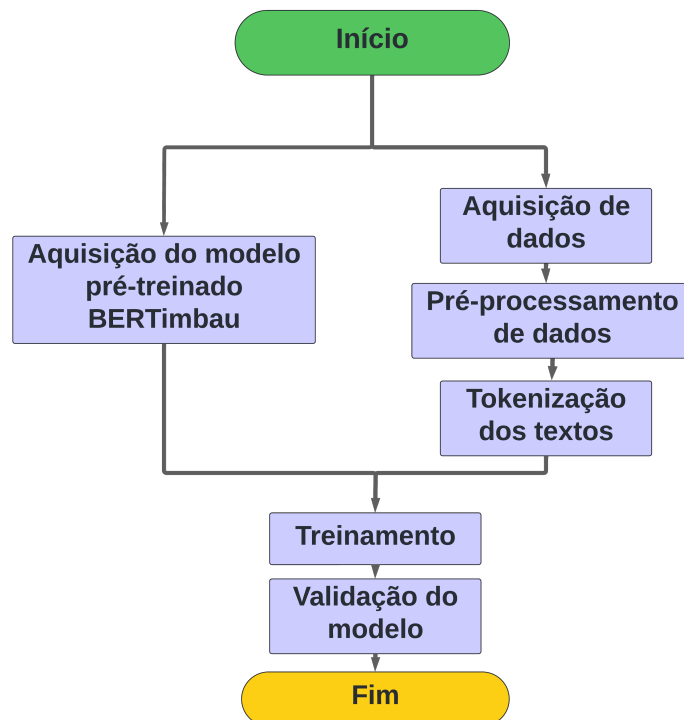
Este comparativo ressalta a diversidade de aplicações e abordagens na área de PLN, destacando a adaptação das metodologias de acordo com o foco específico de cada estudo. O trabalho proposto, por exemplo, segue uma linha similar à de Britto *et al.* (2023) ao se concentrar em um site de vendas, mas se diferencia pelo uso exclusivo do BERTimbau, uma versão do BERT otimizada para o português brasileiro, aplicado à base de dados do Buscapé. A escolha de tecnologias e bases de dados reflete as necessidades específicas de cada investigação, desde a análise de sentimentos até a previsão de comportamento de usuários em diferentes plataformas online.

4 METODOLOGIA

Nesta seção, é detalhado a metodologia utilizada no projeto. Partindo da aquisição do conjunto de dados para análise de sentimentos na língua portuguesa, para isso, foram utilizadas avaliações de usuários do site Buscapé (DAMASCENO, 2023). O primeiro passo foi a limpeza de dados indesejados, sendo retirados todos os textos que se referiam a produtos digitais do conjunto de dados. Também foram realizadas a categorização em quatro formas distintas: na primeira, avaliações com notas 1 e 2 foram classificadas como negativas, nota 3 como neutra, e notas 4 e 5 como positivas; na segunda, notas 1 e 2 foram negativas, enquanto notas 3, 4 e 5 foram positivas; na terceira, classificamos notas 1, 2 e 3 como negativas e 4 e 5 como positivas; na quarta, excluíram-se as avaliações com nota 3, mantendo apenas 1 e 2 como negativas e 4 e 5 como positivas.

Após a organização do conjunto de dados, foi realizada a tokenização dos textos. Seguidamente realizou-se a aquisição do modelo BERTimbau pré-treinado e o ajuste fino com os dados tratados. Por fim, ocorreu-se a validação do modelo. Todo fluxo do processo está ilustrado na Figura 15.

Figura 15 – Fluxo Metodologia.



Fonte: Próprio Autor (2024)

4.1 Base de Dados

A escolha da base de dados de avaliações de usuários do *e-commerce* Buscapé foi motivada principalmente pela característica de cada avaliação estar vinculada a uma classificação numérica de 1 a 5 estrelas, onde há comentários mais diretos, ou mais descritivos, sendo esses apresentados na Tabela 3. Essas classificações refletem o nível de satisfação dos consumidores com os produtos ou serviços avaliados. Esta associação direta entre o texto da avaliação e uma nota quantitativa é extremamente útil, pois oferece *insights* claros sobre os sentimentos e experiências dos usuários de forma.

Tabela 3 – Avaliações de Clientes no Buscapé.

Avaliação (estrelas)	Nome	Data	Comentário
5	Evandro	Janeiro de 2023	Gostei da TV muito bom. Recomendo. Muito bom parabéns.
4	Anderly	Novembro de 2022	Bom. Recomendo.
3	Cleber	Janeiro de 2023	Atualização e funcionamento dos aplicativos pré instalados na tv.
1	Francisco	Dezembro de 2022	A tv apresentou problemas com 15 dias de uso.

Fonte: Buscapé (2024)

Além disso, é importante o fato de essas avaliações estarem redigidas em português-br para a nossa análise. Isso permite uma abordagem mais direcionada e eficaz no contexto brasileiro, facilitando o entendimento das nuances e especificidades da língua. Esta característica linguística é particularmente importante para a análise de sentimentos, onde a precisão na interpretação dos textos é fundamental para resultados confiáveis.

Outro aspecto relevante da escolha desta base de dados é a rica variedade de produtos e serviços avaliados, o que contribui para a diversidade do conjunto de dados. Esta variedade é essencial para garantir que o modelo de análise de sentimentos seja robusto e capaz de interpretar uma ampla gama de contextos e expressões. Adicionalmente, a natureza pública e acessível dessas avaliações oferece uma fonte de dados realista e atualizada, refletindo as tendências e percepções atuais dos consumidores.

4.1.1 Pré-processamento

Esta base de dados é composta pelas seguintes colunas "original_index", "review", "review_processed", "review_tokenized", "polarity", "rating", "kfold_polarity" e "kfold_rating". No entanto, para os propósitos do estudo, focou-se exclusivamente nas colunas "review" e "rating". Essas duas colunas são cruciais para a análise, pois "review" contém o texto das avaliações dos usuários e "rating" representa a classificação atribuída por eles. Esta seleção tem o objetivo de concentrar nossa investigação nos aspectos mais diretamente relacionados às opiniões e avaliações dos usuários, permitindo uma análise mais aprofundada e relevante dentro do contexto específico.

As avaliações dos usuários para produtos são comumente realizadas numa escala de 1 a 5, um método eficaz e simplificado que facilita a expressão do nível de satisfação dos usuários. Esta abordagem permite uma rápida e clara reflexão sobre a opinião dos usuários, com avaliações variando de muito insatisfatórias (pontuação 1) a extremamente positivas (pontuação 5). A ampla gama de classificações oferecidas por essa escala proporciona uma visão abrangente das percepções dos usuários, permitindo a identificação de padrões e tendências nas avaliações. Este sistema de pontuação é um instrumento valioso para avaliar a receptividade dos usuários em relação aos produtos, oferecendo *insights* cruciais para fabricantes e desenvolvedores. A distribuição das avaliações está demonstrada na Tabela 4.

Tabela 4 – Distribuição das Classificações das Avaliações.

Classificação	Quantidade
1.0	3141
2.0	3669
3.0	11364
4.0	33578
5.0	33239

Fonte: Próprio Autor (2024)

O primeiro passo na análise de dados foi a remoção de avaliações de produtos não físicos, como jogos digitais. Esta etapa de limpeza de dados é essencial para garantir que o foco da análise permaneça em produtos tangíveis, cujas avaliações oferecem *insights* diferentes em comparação aos produtos digitais. Após a exclusão desses itens, o próximo passo envolveu um processo de balanceamento do conjunto de dados.

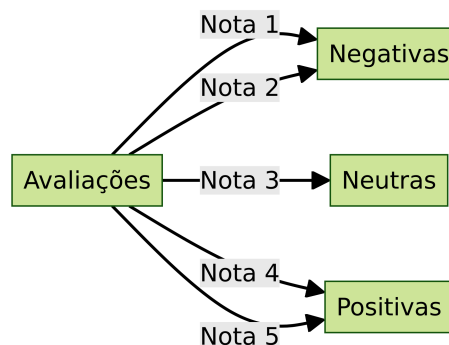
Para alcançar um conjunto de dados equilibrado, foi ajustado a quantidade de comentários para cada categoria de produto, baseando-nos na menor quantidade de avaliações

encontradas em uma categoria individual, que foram de 3.141 avaliações. Este método de balanceamento é essencial para evitar vieses nos resultados da análise, assegurando que nenhuma categoria seja desproporcionalmente representada em comparação com as outras.

O pré-processamento dos dados na análise envolveu um processo de categorização das avaliações, estruturado de quatro maneiras distintas, cada uma visando capturar diferentes nuances nas percepções dos usuários. Esta abordagem diversificada é fundamental para compreender as variadas formas como as avaliações podem ser interpretadas e utilizadas. Vamos detalhar cada uma dessas categorizações:

- **Primeira Categorização - Modelo de Três Classes:** aqui, avaliações com notas 1 e 2 foram classificadas como negativas, a nota 3 como neutra, e notas 4 e 5 como positivas. Esta categorização permite uma diferenciação clara entre experiências negativas, neutras e positivas. A inclusão da categoria neutra é importante, pois oferece um espaço para avaliações que não se encaixam claramente como positivas ou negativas, refletindo as opiniões moderadas ou mistas dos usuários.

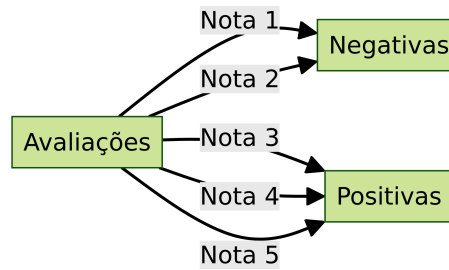
Figura 16 – Modelo de Três Classes.



Fonte: Próprio Autor (2024)

- **Segunda Categorização - Modelo Positivo/Negativo Simplificado:** neste modelo, notas 1 e 2 foram consideradas negativas, enquanto notas 3, 4 e 5 foram agrupadas como positivas. Essa abordagem simplifica a análise ao reduzir as categorias para duas, facilitando a identificação de tendências gerais de satisfação ou insatisfação. Porém, essa simplificação pode mascarar as nuances das avaliações moderadas, que são classificadas como positivas.

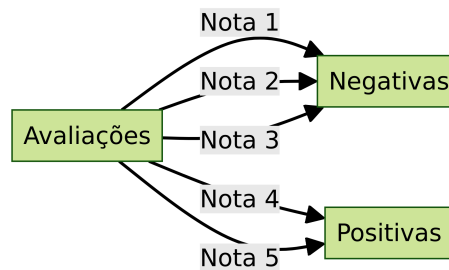
Figura 17 – Modelo Positivo/Negativo Simplificado.



Fonte: Próprio Autor (2024)

- **Terceira Categorização - Modelo Negativo Ampliado:** aqui, foram classificadas as notas 1, 2 e 3 como negativas e 4 e 5 como positivas. Esta categorização é particularmente útil para cenários onde é importante captar uma gama mais ampla de insatisfação do cliente, incluindo avaliações moderadas no espectro negativo, levando a contextos onde até uma experiência moderadamente insatisfatória é significativa para análise e estratégias de melhoria.

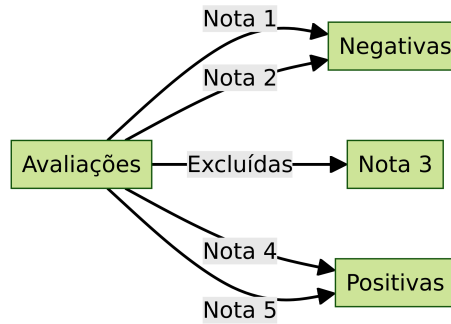
Figura 18 – Modelo Negativo Ampliado.



Fonte: Próprio Autor (2024)

- **Quarta Categorização - Exclusão de Neutros:** nesta abordagem, foram excluídas as avaliações com nota 3, mantendo apenas 1 e 2 como negativas e 4 e 5 como positivas. Este modelo foca exclusivamente nas opiniões polarizadas, removendo as avaliações moderadas da análise. Isso pode ser útil em situações onde o interesse reside estritamente em identificar clientes claramente satisfeitos ou insatisfeitos, descartando avaliações que não representam fortes sentimentos em qualquer direção

Figura 19 – Exclusão de Neutros.



Fonte: Próprio Autor (2024)

4.2 Tokenização

Para que o modelo BERT pudesse processar os textos adequadamente, foi necessário transformar os textos de entrada em uma forma que ele pudesse compreender, o que envolveu a conversão em tokens. Esta etapa foi realizada com o auxílio da biblioteca *Transformers* baseado no trabalho de Devlin *et al.* (2018). A biblioteca não apenas facilitou a tokenização, mas também garantiu que os textos fossem formatados corretamente para o processamento pelo modelo BERT.

Além da tokenização, outro componente utilizado no processamento de textos pelo BERT foram as "máscaras de atenção". Estes eram vetores binários formados por 0s e 1s. As máscaras de atenção auxiliam o BERT a distinguir quais tokens no texto são relevantes para a tarefa específica que estava sendo realizada. Elas sinalizam ao modelo quais partes do texto de entrada devem receber atenção e ser consideradas no processo de análise, enquanto permitem que ele ignore as partes que não são essenciais para o entendimento do contexto ou para a resposta da tarefa específica em questão. Essa capacidade de focar em elementos relevantes e filtrar os irrelevantes foi parte do que tornou o BERT um modelo tão poderoso e eficaz para o processamento de linguagem natural.

4.3 Treinamento

Durante o treinamento dos modelos, optou-se por utilizar os modelos pré-treinados "bert-base-portuguese-cased" (SOUZA *et al.*, 2020). Este modelo específico foi escolhido por sua eficácia e adequação ao processamento de textos em português-br. Para o desenvolvimento do treinamento, segmentamos o conjunto de dados em três partes distintas: 85% foram alocados para o treinamento, enquanto os 7,5% restantes foram divididos igualmente para os conjuntos de

teste e validação.

A plataforma escolhida para realizar o treinamento foi o Google Colab Pro, dada a sua capacidade de processamento e facilidade de uso. As configurações específicas utilizadas estão detalhadas na tabela 5, que inclui informações sobre a alocação de recursos e outras especificações técnicas importantes.

Tabela 5 – Especificações Técnicas do Hardware Utilizado.

Especificações	Estação de Trabalho
CPU	Intel(R) Xeon(R) 2.00GHz
Memória RAM	12.7 GB
Placa De Vídeo	Tesla T4

Fonte: Próprio Autor (2024)

No que diz respeito à configuração dos hiperparâmetros, nos baseamos amplamente no trabalho seminal de Devlin *et al.* (2018), que oferece uma orientação valiosa para o ajuste desses parâmetros em modelos baseados no BERT. Contudo, foi realizado um ajuste no valor do Dropout. Este ajuste foi feito com o objetivo de mitigar o risco de overfitting, ou seja, quando o modelo se ajusta excessivamente aos dados de treinamento a ponto de prejudicar seu desempenho em novos dados. Os valores específicos escolhidos para os hiperparâmetros, incluindo o Dropout, após nossa experimentação, estão registrados detalhadamente em outra tabela, que também pode ser encontrada na tabela 6.

Tabela 6 – Hiperparâmetros Treinamento.

Hiperparâmetro	Utilizado
Optimizer	AdamW
Correct Bias	True
Dropout	2e-3
Learning Rate	2e-5
Max Sequence Length	170
Random Seed	42
Batch Size	16

Fonte: Próprio Autor (2024)

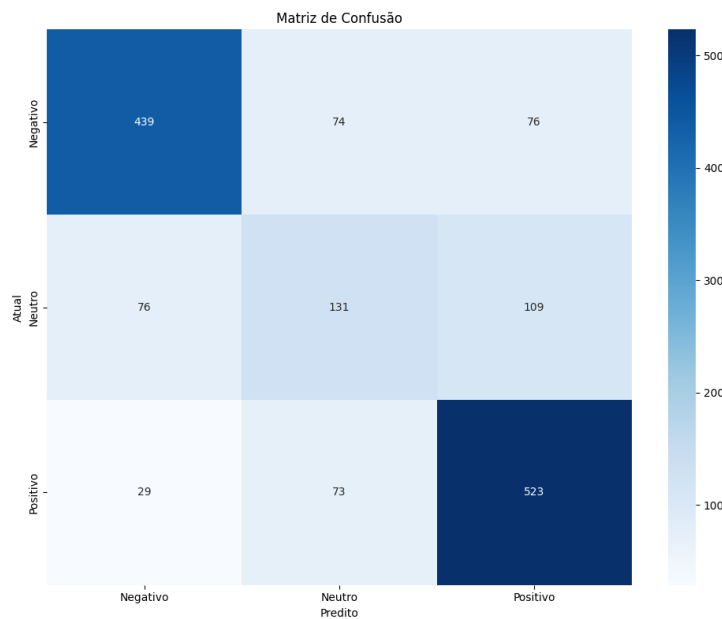
Na primeira fase deste trabalho, foram utilizados para o ajuste fino do modelo os quatro conjuntos de dados de categorização mencionados na subseção 4.1.1. Visando avaliar como o BERTimbau respondeu ao ser treinado com conjuntos de dados variados. Foram observadas métricas como precisão, recall e F1-score para cada arranjo de dataset. Esta análise permite identificar quais configurações de dados produzem os melhores resultados e entender como diferentes tipos de dados influenciam a capacidade de aprendizado do modelo.

5 RESULTADOS

Neste capítulo é apresentado os resultados do treinamento do modelo, apresentados com base na divisão dos arranjos dos datasets, os quais foram aplicados para o fine-tuning (ajuste fino) no modelo BERTimbau visando a análise de sentimentos de avaliações de *e-commerce*.

- **Primeira Configuração:** Esta configuração categoriza as avaliações em negativas, neutras e positivas, atingindo uma acurácia geral de 71%. Observa-se uma precisão de 81% na identificação de sentimentos negativos e de 74% para os positivos. Contudo, a categoria de sentimentos neutros apresenta uma precisão significativamente mais baixa, de apenas 47%. Este desempenho inferior para sentimentos neutros é evidente na matriz de confusão mostrada na Figura 20, onde o modelo demonstra dificuldade em classificar corretamente os valores neutros. Esta dificuldade contribui para a redução da acurácia geral do modelo.

Figura 20 – Primeira Configuração com Valores Negativos, Neutros e Positivos.



Fonte: Próprio Autor (2024)

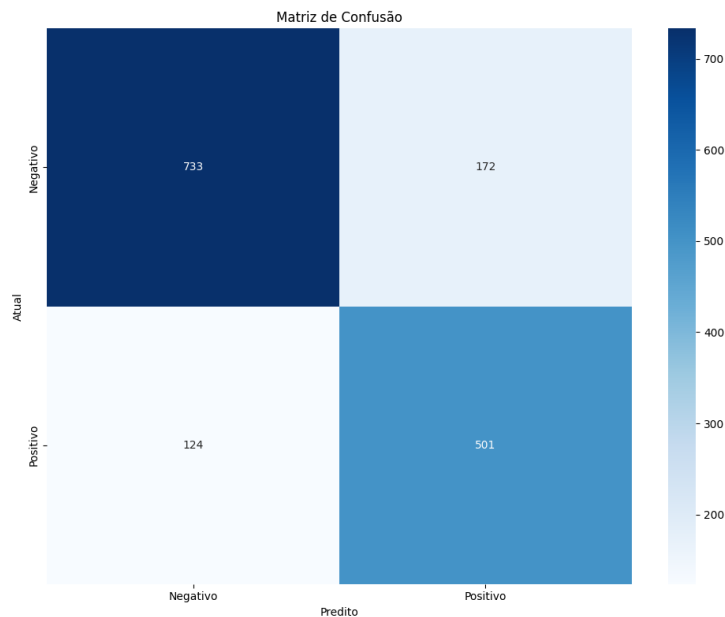
Tabela 7 – Relatório de Classificação da Primeira Configuração.

Classe	Precision	Recall	F1-Score
Negativo	0.81	0.75	0.77
Neutro	0.47	0.41	0.44
Positiva	0.74	0.84	0.78
Accuracy	0.71		

Fonte: Próprio Autor (2024)

- Segunda Configuração:** Aqui, as avaliações são simplificadas em negativas e positivas. O modelo alcançou uma acurácia de 81%. Em termos de precisão, as avaliações negativas foram identificadas com uma precisão de 86%, enquanto as avaliações positivas tiveram uma precisão de 74%. O recall foi de 81% para negativas e 80% para positivas, indicando que o modelo tem uma capacidade similar de identificar corretamente todas as instâncias verdadeiras em ambas as categorias. Já a matriz de confusão apresentada na Figura 21, o modelo teve um desempenho na classificação de avaliações negativas de 733 verdadeiros negativos contra 172 falsos positivos. Por outro lado, houve 124 falsos negativos e 501 verdadeiros positivos, mostrando que, embora o modelo seja confiável na identificação de comentários negativos, ele pode ocasionalmente confundir avaliações negativas com positivas.

Figura 21 – Segunda Configuração com Valores 1, 2 e 3 Negativos e 4 e 5 Positivos.



Fonte: Próprio Autor (2024)

Tabela 8 – Relatório de Classificação Segunda Configuração.

Classe	Precision	Recall	F1-Score
Negativo	0.86	0.81	0.83
Positiva	0.74	0.80	0.77
Accuracy	0.81		

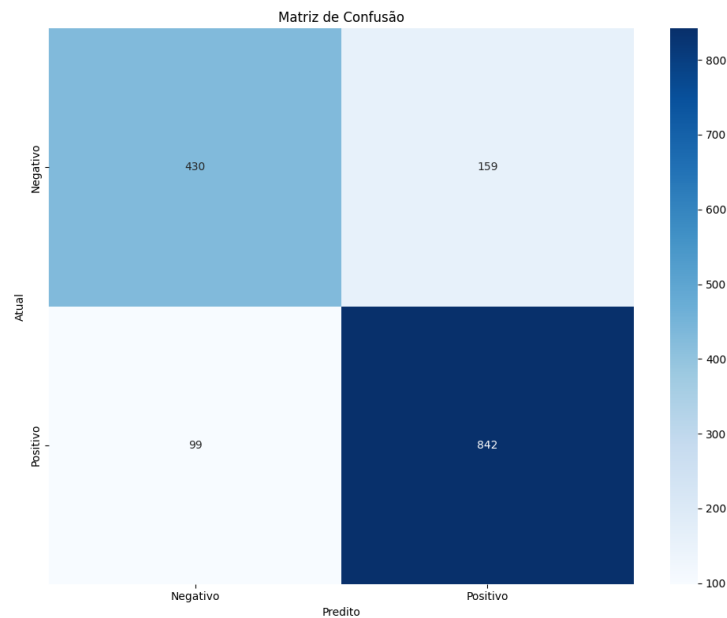
Fonte: Próprio Autor (2024)

- Terceira Configuração:** A acurácia geral alcançada é de 83%. A precisão para as

avaliações negativas é de 81%, enquanto para as positivas é ligeiramente superior, com 84%. O recall para as avaliações negativas é de 73% e para as positivas é notavelmente maior, com 89%. Isso indica que o modelo é mais eficaz em identificar corretamente as avaliações positivas como tal.

A matriz de confusão, ilustrada na Figura 22, mostra que o modelo previu corretamente 430 casos negativos e 842 casos positivos. No entanto, ele incorretamente classificou 159 avaliações negativas como positivas (falsos positivos) e 99 avaliações positivas como negativas (falsos negativos).

Figura 22 – Terceira Configuração com Valores 1 e 2 Negativos e 3, 4 e 5 Positivos.



Fonte: Próprio Autor (2024)

Tabela 9 – Relatório de Classificação da Terceira Configuração.

Classe	Precision	Recall	F1-Score
Negativa	0.81	0.73	0.77
Positiva	0.84	0.89	0.87
Accuracy	0.83		

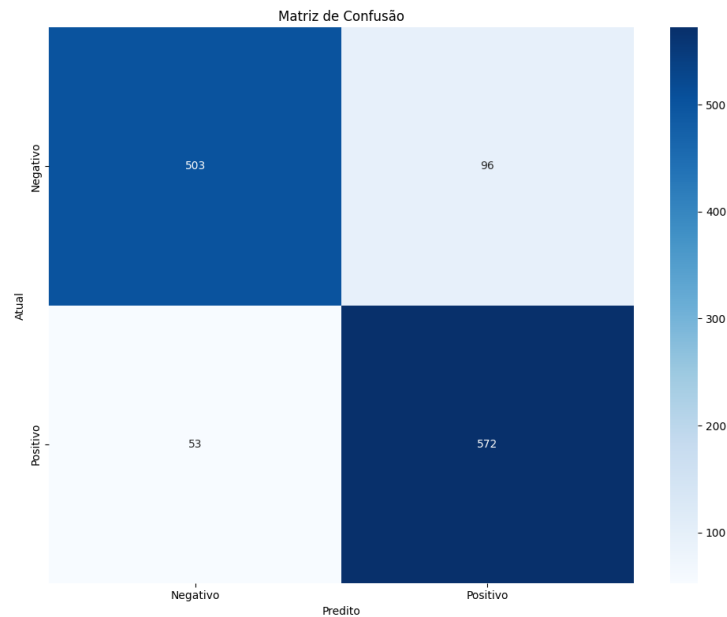
Fonte: Próprio Autor (2024)

- **Quarta Configuração:** Excluindo avaliações neutras, este modelo foca em classificar avaliações claramente positivas ou negativas, tendo uma acurácia de 88%. A precisão para as avaliações negativas é alta, com 90%, enquanto para as positivas é também robusta, com 86%. O recall para as avaliações negativas é de 84% e para as positivas é ainda maior,

com 92%.

A matriz de confusão, ilustrada na Figura 23, mostra que o modelo previu corretamente 503 casos negativos e 572 casos positivos. Contudo, ele incorretamente classificou 96 avaliações negativas como positivas (falsos positivos) e 53 avaliações positivas como negativas (falsos negativos).

Figura 23 – Quarta Configuração com Valores 1 e 2 Negativos, 4 e 5 Positivos e Sem Valores 3.



Fonte: Próprio Autor (2024)

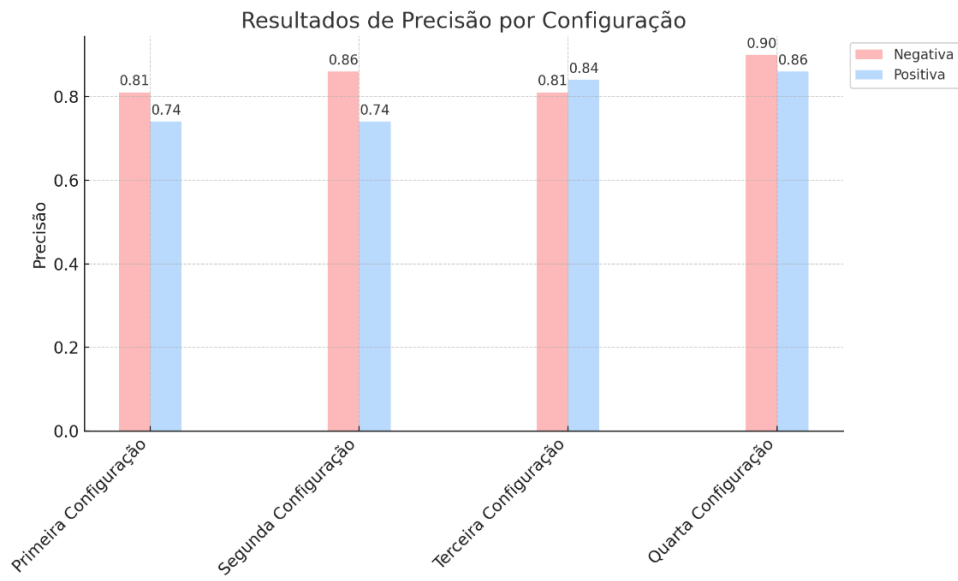
Tabela 10 – Relatório de Classificação da Quarta Configuração

Classe	Precision	Recall	F1-Score
Negativa	0.90	0.84	0.87
Positiva	0.86	0.92	0.88
Accuracy	0.88		

Fonte: Próprio Autor (2024)

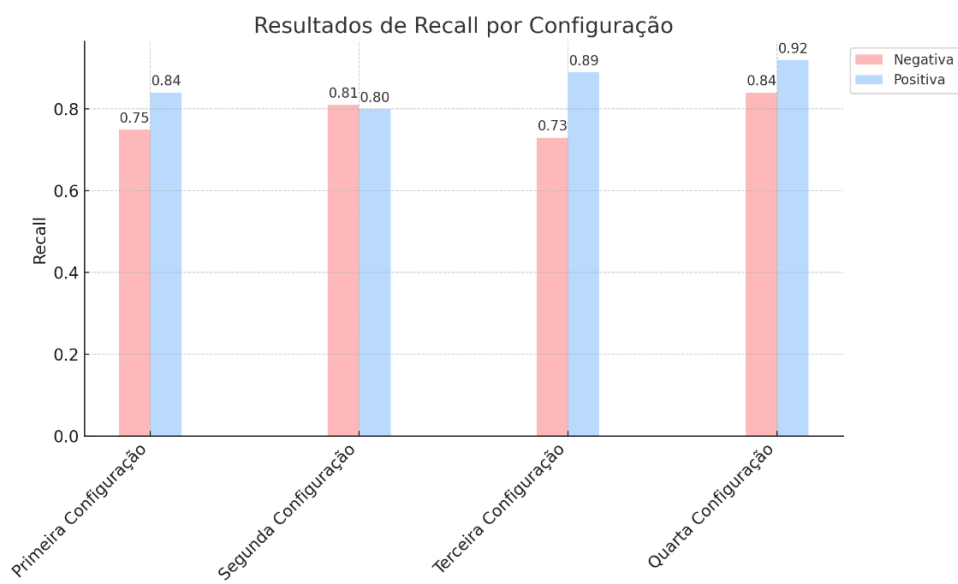
Os gráficos apresentados nas Figuras 24, 25 e 26 destacam os resultados das configurações testadas, evidenciando que a quarta configuração superou as demais em termos de precisão, recall e F1-score. Este resultado sugere que a abordagem de excluir avaliações neutras e focar na classificação direta entre avaliações positivas e negativas proporcionou ao modelo uma capacidade superior de identificação e classificação dos sentimentos expressos nas avaliações de e-commerce.

Figura 24 – Gráfico dos Resultados de Precisão.



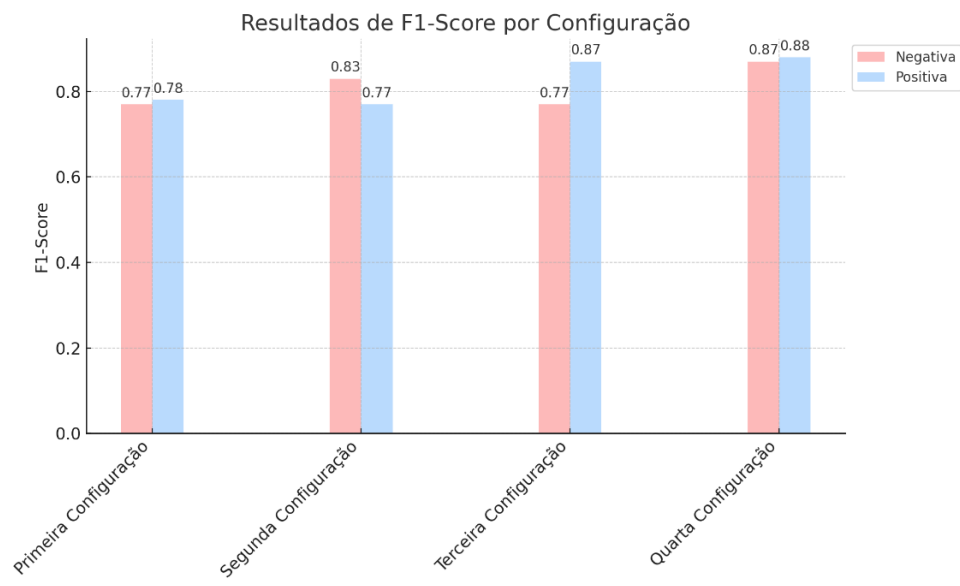
Fonte: Próprio Autor (2024)

Figura 25 – Gráfico dos Resultados de Recall.



Fonte: Próprio Autor (2024)

Figura 26 – Gráfico dos Resultados de F1-score.



Fonte: Próprio Autor (2024)

6 CONCLUSÕES E TRABALHOS FUTUROS

Este estudo ofereceu uma análise detalhada do uso de PLN na avaliação de sentimentos para revisões de *e-commerce* em português, utilizando o modelo BERTimbau. Os resultados demonstraram que técnicas de PLN podem fornecer insights profundos sobre as opiniões dos consumidores brasileiros, contribuindo significativamente para o campo da análise de sentimentos e destacando a eficácia destas técnicas no ambiente de e-commerce.

A Tabela 11 exibe a comparação dos resultados dos trabalhos, onde o trabalho atual obteve uma acurácia de 88%, destacam-se quando comparados com os estudos de Nafees *et al.* (2018) com 71.8% e Britto *et al.* (2023) com 81.3%, mas estão ligeiramente abaixo do alcançado por Lopes *et al.* (2021) onde apresenta 93% de acurácia. Esta diferença pode ser atribuída a vários fatores, incluindo a complexidade das bases de dados, a natureza diversificada das avaliações de produtos em comparação com as de serviços (como acomodações), e as peculiaridades da língua portuguesa refletidas nas avaliações de consumidores brasileiros.

Tabela 11 – Projetos Relacionados a Análise de Sentimentos.

Trabalhos	Nafees <i>et al.</i> (2018)	Britto <i>et al.</i> (2023)	Lopes <i>et al.</i> (2021)	Trabalho Proposto
Foco	Redes Sociais	Site de Vendas	Site de Acomodações	Site de Vendas
Base de Dados	Twitter	Amazon	Airbnb	Buscapé
Tecnologia	NB	SVM	BERTimbau	BERTimbau
Acurácia	71.8%	81.3%	93%	88%

Fonte: Próprio Autor (2024)

Existem diversas oportunidades para pesquisas futuras na área de PLN e análise de sentimentos em língua portuguesa. Algumas sugestões incluem:

- **Expansão Linguística:** Estender a análise para outras variantes da língua portuguesa, como o português europeu, para compreender melhor as diferenças e semelhanças culturais nas avaliações de *e-commerce*.
- **Integração de Fontes de Dados:** Incorporar outras fontes de dados, como redes sociais e fóruns online, para enriquecer a análise de sentimentos e obter uma visão mais holística das percepções do consumidor.

Este estudo reforça a importância da análise de sentimentos para o entendimento do consumidor no e-commerce e sugere caminhos para futuras investigações no campo do PLN aplicado à análise de sentimentos.

REFERÊNCIAS

- AI, D. L. **Natural Language Processing Resources**. 2023. Acessado em: 22 de Janeiro de 2024. Disponível em: <<https://www.deeplearning.ai/resources/natural-language-processing/>>.
- BATES, M.; BOBROW, R. J.; WEISCHEDEL, R. M. In: CAMBRIDGE. **Challenges in Natural Language Processing**. [S.l.]: Cambridge University Press, 1993. (Studies in Natural Language Processing), p. 3–34.
- BRITTO, L. F.; PACÍFICO, L. D.; LUDERMIR, T. B. **Previsão de Utilidade de Avaliações de Produtos Online na Língua Portuguesa Brasileira**. In: SBC. Anais do XIV Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana. [S.l.], 2023. p. 235–239.
- BUSCAPÉ. **Smart TV LED 43” - Buscapé**. 2024. Acessado em 22 de fevereiro de 2024. Disponível em: <<https://www.buscape.com.br/tv/smart-tv-led-43>>.
- CABRAL, A. M.; MARQUES, J. P. C. **How innovation can influence customer satisfaction—case study of the Saccharum Hotel in Madeira**. [S.l.]: Emerald Publishing Limited, 2023. 80–93 p.
- CALIXTO, H. **Tokenização**. 2021. Acessado em 30 de Janeiro de 2024. Disponível em: <<https://linguagemnatural.com.br/significado/tokenizacao/>>.
- CUBE, B. D. **Configurações no Kaggle para o Minicurso de Geoinformática com Aplicações de Machine Learning**. 2020. Acessado em 19 de Janeiro de 2024. Disponível em: <<https://brazil-data-cube.github.io/worcap-2020/getting-started/kaggle.html>>.
- DAMASCENO, F. **Brazilian Portuguese Sentiment Analysis Datasets**. Kaggle, 2023. Acessado em 19 de Janeiro de 2024. Disponível em: <<https://www.kaggle.com/datasets/fredericods/ptbr-sentiment-analysis-datasets>>.
- DENG, L.; LIU, Y. **Deep learning in natural language processing**. [S.l.]: Springer, 2018.
- DEVLIN, J.; CHANG, M.; LEE, K.; TOUTANOVA, K. **BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding**. 2018. Disponível em: <<http://arxiv.org/abs/1810.04805>>.
- GONÇALVES, G. **IAB Brasil revela que 85% dos brasileiros conectados já realizaram compras online**. 2024. Acessado em 27 de Janeiro de 2024. Disponível em: <<https://www.ecommercebrasil.com.br/noticias/iab-brasil-revela-que-85-dos-brasileiros-conectados-ja-realizaram-compras-online>>.
- HANSEL, L. **Como as avaliações online influenciam as decisões de compra dos consumidores**. 2024. Acessado em: 22 de Janeiro de 2024. Disponível em: <<https://harmo.me/blog/avaliacoes-online-influencias-decisoes>>.
- HAPKE, H.; HOWARD, C.; LANE, H. **Natural Language Processing in Action: Understanding, analyzing, and generating text with Python**. [S.l.]: Simon and Schuster, 2019.
- Hugging Face. **Hugging Face Organization Profile on Qiita**. 2024. Acessado em 21 de Fevereiro de 2024. Disponível em: <<https://qiita.com/organizations/huggingface>>.

- IAB Brasil; Offerwise. **Comportamento de compra online: hábitos de compra dos brasileiros na internet e sua relação com anúncios**. 2023. Acessado em 20 de fevereiro de 2024. Disponível em: <<https://iabbrasil.com.br/pesquisa-comportamento-de-compra-online/>>.
- IACOBUCCI, D.; PETRESCU, M.; KRISHEN, A.; BENDIXEN, M. **The state of marketing analytics in research and practice**. [S.l.]: Springer, 2019. 152–181 p.
- JORDAN, M. I.; MITCHELL, T. M. **Machine learning: Trends, perspectives, and prospects**. [S.l.]: American Association for the Advancement of Science, 2015. 255–260 p.
- JURAFSKY, D.; MARTIN, J. H. **Speech and Language Processing**. [S.l.]: Pearson Prentice Hall, 2009.
- KAPLAN, A.; HAENLEIN, M. **Siri, Siri, in my hand: Who’s the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence**. [S.l.]: Elsevier, 2019. 15–25 p.
- KINAST, P. **Alexa, Siri ou Google Assistente: Qual é o melhor?** 2023. Acessado em 30 de Janeiro de 2024. Disponível em: <<https://www.oficinadanet.com.br/tecnologia/28456-alexa-siri-ou-google-assistente>>.
- LEE, P. S.; CHAKRABORTY, I.; BANERJEE, S. **Artificial Intelligence applications to customer feedback research: A review**. [S.l.]: Emerald Publishing Limited, 2023. 169–190 p.
- LEINER, B. M.; CERF, V. G.; CLARK, D. D.; KAHN, R. E.; KLEINROCK, L.; LYNCH, D. C.; POSTEL, J.; ROBERTS, L. G.; WOLFF, S. **A brief history of the Internet**. [S.l.]: ACM New York, NY, USA, 2009. 22–31 p.
- LIU, B. **Sentiment analysis and opinion mining**. [S.l.]: Springer Nature, 2022.
- LOPES, E.; CORREA, U.; FREITAS, L. **Exploring bert for aspect extraction in portuguese language**. 2021.
- MANNING, C.; SCHUTZE, H. **Foundations of statistical natural language processing**. [S.l.]: MIT press, 1999.
- MEHTA, R. **Natural Language Processing (NLP) [A Complete Guide]**. 2023. Acessado em: 22 de Janeiro de 2024. Disponível em: <<https://www.deeplearning.ai/resources/natural-language-processing/>>.
- MindSpore. **Using the BERT Network to Implement Intelligent Poem Writing**. 2022. Acessado em 02 de Fervereiro de 2024. Disponível em: <https://www.mindspore.cn/tutorial/training/en/r1.2/advanced_use/nlp_bert_poetry.html>.
- Ministério das Comunicações. **80% dos domicílios brasileiros possuem acesso à internet, aponta pesquisa**. 2023. Acessado em 27 de Janeiro de 2024. Disponível em: <<https://www.gov.br/mcom/pt-br/noticias/2023/maio/80-dos-domicilios-brasileiros-possuem-acesso-a-internet-aponta-pesquisa>>.
- NAFEES, M.; DAR, H.; LALI, I. U.; TIWANA, S. **Sentiment analysis of polarity in product reviews in social media**. 2018. 1–6 p.

- NOGARE, D. **Performance de Machine Learning - Matriz de Confusão**. 2020. Acessado em 02 de Fevereiro de 2024. Disponível em: <<https://diegonogare.net/2020/04/performance-de-machine-learning-matriz-de-confusao/>>.
- PANDEY, S. V.; DEORANKAR, A. **A study of sentiment analysis task and it's challenges**. 2019. 1–5 p.
- RUSSELL, S. J.; NORVIG, P. **Artificial intelligence a modern approach**. [S.l.]: London, 2010.
- SANH, V.; DEBUT, L.; WOLF, T. **DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter**. 2019.
- SHIMABUKURO, I. **Segmento de e-commerce pode crescer 56% no Brasil até 2024**. 2021. Acessado em 24 de Janeiro de 2024. Disponível em: <<https://olhardigital.com.br/2021/03/24/pro/segmento-de-e-commerce-pode-crescer-cinquenta-e-seis-por-cento-no-brasil-ate-2024/>>.
- SISTA, R.; SINGH, R.; KUMAWAT, S. K.; DHANARE, R. **Techniques used by E-commerce industries for Customer analysis**. 2021. 1-6 p.
- SOUZA, F.; NOGUEIRA, R.; LOTUFO, R. **BERTimbau: pretrained BERT models for Brazilian Portuguese**. 2020. 403–417 p.
- STATISTA. **Online Shopping - Statistics Facts**. 2020. Acessado em: 22 de Janeiro de 2024. Disponível em: <<https://www.statista.com/topics/871/online-shopping/#topicOverview>>.
- SUPEROFFICE. **Customer Experience Strategy**. 2024. Acesso em 29 de janeiro de 2024, às 18:44 -04. Disponível em: <<https://www.superoffice.com/blog/customer-experience-strategy/>>.
- SYAM, N.; SHARMA, A. **Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice**. [S.l.]: Elsevier, 2018. 135–146 p.
- VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, L.; POLOSUKHIN, I. **Attention Is All You Need**. 2017.
- VERHOEF, P. C.; LEMON, K. N.; PARASURAMAN, A.; ROGGEVEEN, A.; TSIROS, M.; SCHLESINGER, L. A. **Customer experience creation: Determinants, dynamics and management strategies**. [S.l.]: Elsevier, 2009. 31–41 p.
- WANDERLEY, C. A. C.; COSTA, R. S. da; RIBEIRO, L. de P. **Crimes Cibernéticos Em Tempos De Pandemia: O Isolamento Social Como Propulsor Da Vulnerabilidade Da População E Do Aumento Dos Casos**. 2022.
- ZHANG, B.; XU, X.; LI, X.; CHEN, X.; YE, Y.; WANG, Z. **Sentiment analysis through critic learning for optimizing convolutional neural networks with rules**. 2019. 21-30 p. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0925231219306198>>.