

UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

**Gabriel Barbosa Cândido**

**Análise de Sentimentos no Domínio do  
E-commerce Utilizando Modelos de  
Linguagem**

São Carlos - SP  
2024



**Gabriel Barbosa Cândido**

**Análise de Sentimentos no Domínio do  
E-commerce Utilizando Modelos de  
Linguagem**

Trabalho de Conclusão de Curso apresentado ao Departamento de Engenharia Elétrica do Centro de Ciências Exatas e de Tecnologia da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Bacharel em Engenharia Elétrica.

Área de concentração: Inteligência Artificial

Orientador: Prof. Dra. Helena de Medeiros Caseli

São Carlos - SP  
2024



*Dedico este trabalho a Deus, fonte inesgotável de sabedoria e guia constante em minha jornada acadêmica. A Ele, agradeço por iluminar meu caminho, fortalecer minha determinação e conceder-me a graça de concluir este desafio. Dedico também este esforço aos meus familiares e amigos, cujo apoio e compreensão foram fundamentais. Este TCC é uma expressão de gratidão a todos que contribuíram para o meu crescimento acadêmico e pessoal.*



---

# Agradecimentos

---

Primeiramente, expresso minha imensa gratidão a Deus por guiar-me ao longo de toda a jornada acadêmica; a Ele, sou profundamente grato.

À minha futura esposa, Nathália de Souza Thomazini, constante fonte de amor, apoio e inspiração. Sua presença ilumina minha vida, tornando esta jornada muito mais significativa.

Aos meus amados pais, Rogério Domingos Cândido e Mércia Peris Barbosa Cândido, e à minha querida avó, Neide Peris Barbosa, cujo amor incondicional e sabedoria moldaram a pessoa que sou hoje. Suas bênçãos e encorajamentos foram alicerces essenciais nesta jornada.

Aos meus sogros, Washington Luiz Thomazini e Maria José de Souza, por acolherem-me na família com tanto amor e gentileza.

À professora Helena de Medeiros Caseli, cuja orientação e sabedoria foram fundamentais para o desenvolvimento deste trabalho.

Por fim, aos meus queridos amigos Rafael Ramos Teles e Pedro Henrique Bogaz Manfrim, cuja amizade e apoio foram pilares essenciais durante esta jornada. Compartilhar este caminho ao lado de amigos tão especiais enriqueceu toda minha trajetória acadêmica.



*“A humildade é o primeiro degrau para a sabedoria”  
(Santo Tomás de Aquino)*



---

# Resumo

---

O processamento de linguagem natural (PLN) é um ramo da ciência da computação, mais especificamente da Inteligência Artificial, cujo o objetivo é dar aos computadores a capacidade de compreender palavras, tanto em formato de texto como de voz, da mesma forma que nós, seres humanos, as conseguimos compreender. Um dos domínios onde o PLN vem sendo amplamente utilizado é o do e-commerce (comércio eletrônico). Quando os consumidores fazem uma compra, eles geralmente deixam as suas opiniões a respeito de um determinado produto e, através do conteúdo emocional dessas opiniões, as empresas conseguem identificar as áreas que necessitam de melhoria. Dentro do escopo de PLN, a tarefa de Análise de Sentimentos visa resolver esse tipo de problema através da classificação do conteúdo emocional por detrás dessas opiniões. Nesse contexto, este trabalho tem como principal objetivo comparar diferentes abordagens de redes neurais baseadas em arquitetura de *transformers*. Dentre esses modelos, destaca-se o BERT (Bidirectional Encoder Representations from Transformers) com uma notável evolução na arquitetura de transformadores. O BERT, juntamente com suas variantes, tem demonstrado proeminência como uma classe de modelos que superam o processamento sequencial de informações. Dentre as variantes do BERT, o BERTimbau, treinado especificamente para o Português brasileiro, destaca-se como um modelo de referência em tarefas como similaridade semântica, inferência textual e reconhecimento de entidades nomeadas. Este trabalho realiza o *fine-tuning* dos modelos BERT, RoBERTa e BERTimbau na tarefa de classificação de sentimentos, utilizando o *corpus* B2W-Reviews. O objetivo é estabelecer um *benchmark* dos modelos no contexto da Análise de Sentimentos para o idioma português.

**Palavras-chave:** E-commerce. Análise de Sentimentos. Transformadores .



---

# Abstract

---

Natural Language Processing (NLP) is a subfield of computer science, specifically within Artificial Intelligence, aimed at providing computers with the ability to comprehend words, both in text and speech, similar to how humans do. One domain where NLP has been extensively applied is e-commerce. When consumers make purchases, they often share opinions about products, enabling companies to identify areas for improvement through the emotional content of these reviews. Within the scope of NLP, the task of Sentiment Analysis aims to address this issue by classifying the emotional content behind these opinions. In this context, the primary objective of this work is to compare different approaches of neural networks based on transformer architectures. Notably, the BERT (Bidirectional Encoder Representations from Transformers) model stands out with a remarkable evolution in transformer architecture. BERT, along with its variants, has demonstrated prominence as a class of models that surpass sequential information processing. Among the variants of BERT, BERTimbau, specifically trained for Brazilian Portuguese, stands out as a benchmark model in tasks such as semantic similarity, textual inference, and named entity recognition. This study conducts fine-tuning of the BERT, RoBERTa, and BERTimbau models for sentiment classification, utilizing the B2W-Reviews corpus. The goal is to establish a benchmark in the context of Sentiment Analysis for the Portuguese language.

**Keywords:** E-commerce. Sentiment Analysis. Transformers. .



---

# Lista de ilustrações

---

Figura 1 – Geração das word embeddings . . . . .	27
Figura 2 – Representação das palavras com codificação One-Hot. . . . .	27
Figura 3 – Representação espacial das word embeddings . . . . .	28
Figura 4 – Arquitetura Transformers Simplificada . . . . .	29
Figura 5 – Ilustração do mecanismo de self-attention . . . . .	30
Figura 6 – Geração dos vetores Q e (K-V). . . . .	31
Figura 7 – Cálculo do mecanismo de atenção . . . . .	31
Figura 8 – Arquitetura de Transformers em Alto Nível . . . . .	32
Figura 9 – Representação das Entradas das Sentenças no BERT. . . . .	34
Figura 10 – Mascaramento Estático x Dinâmico. . . . .	36
Figura 11 – Comparativo entre os modelos para inglês e português. . . . .	40
Figura 12 – Comparativo entre os modelos para inglês e português. . . . .	41
Figura 13 – Distribuição de Polaridades . . . . .	45
Figura 14 – Distribuição da densidade de tokens no B2W-Reviews . . . . .	46
Figura 15 – Comparativo entre os modelos para inglês e português. . . . .	49
Figura 16 – Matriz de Confusão do Modelo BERT . . . . .	52
Figura 17 – Matriz de Confusão do Modelo RoBERTa . . . . .	52
Figura 18 – Matriz de Confusão do Modelo BERTimbau . . . . .	53



---

# Lista de tabelas

---

Tabela 1 – Valores dos Hiperparâmetros . . . . .	47
Tabela 2 – Métricas de Classificação para Cada Modelo . . . . .	51
Tabela 3 – Exemplos Errôneos na Rotulagem do Corpus B2W-Reviews . . . . .	54
Tabela 4 – Exemplos de Erros Cometidos pelo Modelo ao Prever a Classe como Neutra . . . . .	55



---

# Lista de siglas

---

**BERT** Bidirectional Encoder Representations from Transformers

**RoBERTa** A Robustly Optimized BERT Pretraining Approach

**AS** Análise de Sentimentos

**NER** Reconhecimento de Entidades Nomeadas

**STS** Similaridade Textual de Sentenças

**CGU** Conteúdo Gerado por Usuário

**CB** Class Balanced Loss



---

# Sumário

---

<b>1</b>	<b>INTRODUÇÃO . . . . .</b>	<b>21</b>
1.1	Objetivo Geral . . . . .	22
1.2	Objetivos Específicos . . . . .	22
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA . . . . .</b>	<b>23</b>
2.1	Análise de Sentimentos . . . . .	23
2.1.1	Diferentes níveis de análise . . . . .	23
2.2	O Normalizador Enelvo . . . . .	24
2.2.1	Módulo de Pré-processamento . . . . .	25
2.2.2	Módulo de Análise . . . . .	25
2.2.3	Módulo de Geração de Candidatos . . . . .	26
2.2.4	Módulo de Ranqueamento . . . . .	26
2.3	Word Embeddings . . . . .	26
2.3.1	Positional Encoder . . . . .	28
2.4	Transformers . . . . .	29
2.4.1	Attention . . . . .	30
2.4.2	Encoder-Decoder . . . . .	32
2.5	BERT . . . . .	33
2.5.1	Arquitetura do Modelo . . . . .	34
2.5.2	Modelagem de Linguagem Mascarada . . . . .	35
2.6	RoBERTa . . . . .	35
2.6.1	Dynamic Masking . . . . .	35
2.6.2	Otimizações . . . . .	36
2.6.3	BERTimbau . . . . .	37
<b>3</b>	<b>TRABALHOS RELACIONADOS . . . . .</b>	<b>39</b>

<b>3.1</b>	<b>Utilizando BERTimbau para a Classificação de Emoções em Português . . . . .</b>	<b>39</b>
<b>3.2</b>	<b>Fine-grained Sentiment Classification using BERT . . . . .</b>	<b>40</b>
<b>3.3</b>	<b>Comparando os Trabalhos . . . . .</b>	<b>41</b>
<b>4</b>	<b>DESENVOLVIMENTO . . . . .</b>	<b>43</b>
<b>4.1</b>	<b>Ferramentas . . . . .</b>	<b>43</b>
4.1.1	Tratamento de Dados . . . . .	43
4.1.2	Criação de Modelos . . . . .	44
<b>4.2</b>	<b>Base de Dados . . . . .</b>	<b>44</b>
<b>4.3</b>	<b>Pré-Processamento . . . . .</b>	<b>45</b>
<b>4.4</b>	<b>Modelos Base . . . . .</b>	<b>45</b>
4.4.1	BERT e BERTimbau . . . . .	45
4.4.2	RoBERTa . . . . .	46
<b>4.5</b>	<b>Treino . . . . .</b>	<b>47</b>
4.5.1	Dados de Entrada . . . . .	47
4.5.2	Loss Function . . . . .	47
4.5.3	Hiperparâmetros . . . . .	47
<b>4.6</b>	<b>Métricas de Avaliação . . . . .</b>	<b>48</b>
<b>5</b>	<b>EXPERIMENTOS E RESULTADOS . . . . .</b>	<b>51</b>
<b>5.1</b>	<b>Métricas dos Modelos . . . . .</b>	<b>51</b>
<b>5.2</b>	<b>Matrizes de Confusão . . . . .</b>	<b>51</b>
<b>5.3</b>	<b>Exemplos de Classificação . . . . .</b>	<b>53</b>
	<b>Referências . . . . .</b>	<b>59</b>

---

# Capítulo 1

## Introdução

---

Nos últimos anos, o cenário dos e-commerce tem testemunhado um notável aumento no volume de avaliações e opiniões publicadas em suas plataformas. Os consumidores, por meio de um simples comentário, agora detêm o poder de popularizar ou criticar um produto ou serviço. Essa capacidade de expressar opiniões e feedback tem se tornado uma fonte valiosa de informações para empresas, permitindo aprimorar seus produtos e serviços para atender às expectativas de seus potenciais clientes.

Paralelamente, no âmbito da pesquisa acadêmica, a crescente quantidade de informações não estruturadas, provenientes de avaliações e opiniões, é objeto de estudo de uma área conhecida como Análise de Sentimentos (AS). De acordo com a definição de (Pang e Lee 2008), a AS engloba a “tarefa de determinar a polaridade expressa em um documento de texto”, onde um “documento” pode ser uma sentença, um parágrafo ou até mesmo um texto completo. A Análise de sentimentos se beneficia da convergência entre a mineração de dados e o processamento de linguagem natural. Seu objetivo é não apenas compreender os sentimentos expressos nas opiniões, mas também classificá-los de maneira eficiente, permitindo a extração de insights valiosos a partir de uma variedade de fontes textuais não estruturadas.

Uma das questões centrais da AS e, de particular interesse para esta pesquisa, é a seleção do modelo de linguagem natural mais eficaz para a análise de sentimentos no contexto dos e-commerce. Conforme destacado por (Liu 2012), “a escolha da abordagem correta em análise de sentimentos é crucial para a precisão dos resultados.” Com a disponibilidade de diversas técnicas e algoritmos, a seleção da abordagem ideal é de extrema importância, uma vez que impacta diretamente na qualidade da análise de sentimentos e, conseqüentemente, nos resultados das decisões tomadas a partir dela.

Neste contexto, uma das mais recentes inovações, foi o surgimento da arquitetura

transformers. Essa arquitetura possibilitou que o treinamento dos modelos fosse realizado em grandes corpora para uma tarefa geral de Processamento de Linguagem Natural (PLN) e que pudessem ser ajustados finamente (*fine-tuning*) para outras tarefas de PLN, como a AS. Uma das grandes vantagens desse tipo de arquitetura é que elas conseguem representar de forma mais precisa as palavras contidas nos textos, capturando adequadamente o contexto em que as palavras estão inseridas. Para este trabalho, serão exploradas abordagens baseadas em transformers, com foco nos modelos BERT e suas variantes, como o RoBERTa e o BERTimbau, realizando a adaptação dos modelos para tarefa de classificação de sentimentos no domínio dos e-commerce.

## 1.1 Objetivo Geral

O objetivo geral deste trabalho é investigar a aplicação de diversos modelos de linguagem na tarefa de Análise de Sentimentos (AS) no contexto dos e-commerce, a fim de determinar qual deles é o mais eficaz para essa tarefa. Isso será alcançado por meio do *fine-tuning* de modelos de linguagem estado-da-arte e da avaliação de seu desempenho em relação ao benchmark estabelecido com o uso do dataset B2W-Reviews.

## 1.2 Objetivos Específicos

- ❑ Estudar as abordagens de Análise de Sentimentos existentes na literatura acadêmica e no contexto do comércio eletrônico.
- ❑ Coletar, pré-processar e analisar o dataset B2W-Reviews para criar um conjunto de dados de avaliações de e-commerce útil para esta pesquisa.
- ❑ Implementar e comparar o desempenho de diversos modelos de linguagem natural, com destaque para abordagens baseadas em transformers, principalmente o BERT e suas variantes.
- ❑ Avaliar o desempenho dos modelos em termos de precisão, recall, f1-score e outras métricas relevantes.
- ❑ Criar um benchmark para a análise de sentimentos no domínio dos e-commerce com base nos resultados obtidos.

---

# Capítulo 2

## Fundamentação teórica

---

Neste capítulo serão apresentados: o conceito de análise de sentimentos e os seus diferentes níveis de análise, os métodos empregados nas fases de pré-processamento das sentenças e o funcionamento dos modelos de linguagem natural usados na tarefa de classificação de polaridade.

### 2.1 Análise de Sentimentos

A análise de sentimentos, também conhecida como mineração de opinião, é um sub-campo que emprega técnicas computacionais para classificar emoções, opiniões e atitudes expressas em textos. Por meio da utilização de técnicas de PLN combinadas com métodos de aprendizado de máquina, a análise de sentimentos busca identificar o teor emocional presente nos documentos de textos e classificá-los como positivo, negativo ou neutro.

#### 2.1.1 Diferentes níveis de análise

De acordo com (Liu 2012), a análise de sentimentos tem sido investigada predominantemente em três níveis de granularidade:

- ❑ **Nível de documento:** Este nível envolve a avaliação da polaridade geral expressa em um documento completo. Por exemplo, ao analisar uma review de produto, a atribuição de polaridade considerará a totalidade da review como um único documento. Isso implica a busca por um sentimento global positivo, negativo ou neutro no conteúdo da review como um todo.

- **A nível de sentença:** Este nível aprofunda-se nas sentenças, avaliando se cada uma delas expressa uma opinião positiva, negativa ou neutra. No contexto, “neutro”, indica um sentimento de indiferença. Essa análise está estreitamente vinculada à classificação de subjetividade (Wiebe Janyce M. e O’Hara 1999), que diferencia as sentenças expressando informações factuais denominadas objetivas daquelas que refletem pontos de vista e opiniões pessoais denominadas subjetivas. Contudo, é crucial notar que subjetividade não se confunde com sentimento, visto que muitas sentenças objetivas podem sutilmente implicar opiniões, como no exemplo dado por (Liu 2012): “Compramos o carro no mês passado e o limpador de para-brisa caiu.”
  
- **Nível de entidade e aspecto:** Tanto as análises em nível de documento quanto em nível de sentença não revelam especificamente o que exatamente as pessoas gostaram ou não gostaram. O nível de aspecto realiza uma análise mais refinada, esse nível não se baseia em construtos linguísticos convencionais (documentos, parágrafos, sentenças, cláusulas ou frases). Em vez disso, foca diretamente na opinião em si. Essa abordagem se apoia na ideia de que uma opinião é composta por um sentimento (positivo ou negativo) e um alvo (do opinador). Uma opinião sem a identificação de seu alvo possui utilidade limitada. Reconhecer a importância dos alvos de opinião também contribui para uma compreensão mais aprofundada do problema de análise de sentimentos. Em (Liu 2012), o autor exemplifica com a frase “embora o serviço não seja ótimo, eu ainda amo este restaurante” que embora tenha um tom positivo, não podemos afirmar que a sentença seja totalmente positiva. Na verdade, a sentença é positiva sobre o restaurante (ênfático), mas negativa sobre o serviço (não ênfático). Em muitas aplicações, os alvos de opinião são descritos por entidades e/ou seus diferentes aspectos. Assim, o objetivo deste nível de análise é descobrir sentimentos sobre entidades e/ou seus aspectos.

## 2.2 O Normalizador Enelvo

A normalização de textos desempenha um papel crucial na aprimoramento da compreensão do Conteúdo Gerado por Usuário (CGU). No contexto específico da análise de sentimentos em reviews de produtos, a relevância desse processo torna-se ainda mais evidente. Ao nos depararmos com textos oriundos de consumidores, é comum encontrarmos discrepâncias e imprecisões que, infelizmente, podem prejudicar a clareza e a coesão do conteúdo.

Nesse cenário, destaca-se o Enelvo como uma ferramenta especializada dedicada à normalização do Conteúdo Gerado por Usuário na língua portuguesa. O propósito fundamental do Enelvo é identificar e corrigir possíveis ruídos ou anomalias presentes nos textos, com o intuito de aprimorar a qualidade textual. Sua atuação como um meca-

nismo de aprimoramento textual vai além da simples correção de erros gramaticais. Ao contrário, o Enelvo contempla também aspectos específicos da linguagem empregada em contextos online, como emojis, hashtags e nomes de usuários de Twitter.

Os autores (Bertaglia e Nunes 2016) resumem o funcionamento da ferramenta da seguinte maneira.

1. **Entrada da Sentença:** A ferramenta recebe como entrada uma sentença de texto gerada por um usuário.
2. **Módulo de Pré-processamento:** Em seguida, a sentença é pré-processada com a finalidade de deixar no formato correto utilizado pelo próximo módulo.
3. **Módulo de Análise:** Após esse processo, o módulo de análise é responsável por identificar possíveis ruídos na sentença.
4. **Módulo de Geração de Candidatos:** Esses ruídos são recebidos pelo módulo de geração de candidatos fazendo possíveis correções no texto.
5. **Módulo de Ranqueamento:** Neste módulo acontece a atribuição de pontuação para cada um dos candidatos, possibilitando a substituição do ruído por uma palavra adequada.
6. **Sentença Normalizada:** Por fim, a sentença é retornada corrigida, substituindo os ruídos que foram encontrados por palavras coerentes.

### 2.2.1 Módulo de Pré-processamento

No módulo de pré-processamento acontece a segmentação de sentenças, inserção de pontuação, capitalização, tokenização e remoção de repetições excessivas. A segmentação de sentenças é realizada por um sistema já consolidado, proposto por Condori e Pardo (2015), porém, o normalizador se destaca pela capacidade de identificar emojis, hashtags e nomes de usuários de Twitter. Além disso, ele permite agrupar entidades relacionadas a números como tokens únicos e oferece opções de substituição de determinados tipos de tokens. A saída do pré-processamento é um texto segmentado em sentenças, tokenizado e capitalizado de acordo com a pontuação inserida pelo segmentador, pronto para análises subsequentes na CGU.

### 2.2.2 Módulo de Análise

O módulo de análise assume uma função essencialmente quantitativa ao produzir estatísticas descritivas relativas às amostras de texto sob análise. Esse enfoque estatístico visa facilitar a identificação de ruídos no processo, tornando-o mais intuitivo. Alguns exemplos de estatística, como mencionado pelos autores (Bertaglia e Nunes 2016) são as

palavras discrepantes de um padrão e a média de tais palavras por sentença. Essas métricas proporcionam uma visão quantitativa e esclarecedora das características textuais, fundamentais para aprimorar a compreensão e a eficácia do processo de análise.

### 2.2.3 Módulo de Geração de Candidatos

O módulo em destaque desempenha uma função essencial ao analisar palavras que se desviam do padrão estabelecido, propondo candidatos para sua normalização. Neste contexto, são apresentados os métodos de geração por similaridade lexical e através de word embeddings. No que concerne à abordagem de similaridade lexical, o sistema oferece flexibilidade ao permitir a aplicação de diversas medidas, não restritas às consideradas neste estudo. Adicionalmente, é viável combinar diferentes medidas para gerar candidatos, proporcionando, assim, uma abordagem mais abrangente.

### 2.2.4 Módulo de Ranqueamento

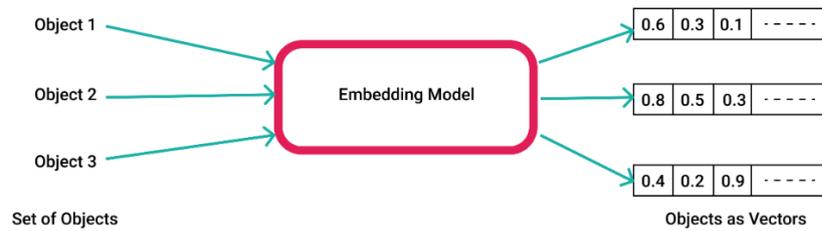
Na fase subsequente à geração de múltiplos candidatos para a normalização de cada palavra, emerge a necessidade de uma avaliação criteriosa para determinar a opção mais apropriada para efetuar a substituição. Este módulo desempenha um papel crítico nesse procedimento, empregando técnicas fundamentadas em similaridade lexical e word embeddings, de forma análoga à etapa de geração de candidatos. A saída resultante deste módulo pode se apresentar como o candidato de maior pontuação ou, alternativamente, como uma lista ordenada dos  $n$  melhores candidatos, sendo  $n$  um parâmetro variável definido pela aplicação.

## 2.3 Word Embeddings

O conceito de word embeddings reside na representação de todas as palavras por meio de vetores densos de números reais. Esses vetores são aprendidos para representar o texto em um espaço de múltiplas dimensões, onde palavras com significados similares possuem representações semelhantes. Essa abordagem implica que palavras relacionadas estão próximas no espaço vetorial de palavras. Portanto, ao construir um espaço de word embeddings, o objetivo fundamental é capturar diferentes tipos de relações nesse espaço, que podem ser associadas ao significado, morfologia, contexto ou qualquer outro tipo de relação intrínseca entre as palavras.

A Figura 1 ilustra a geração de vetores a partir de objetos de entrada. Essa abordagem é amplamente utilizada, visto que as representações de significado das palavras podem ser aplicadas em diversas tarefas, como análise de sentimentos ou tradução automática de textos.

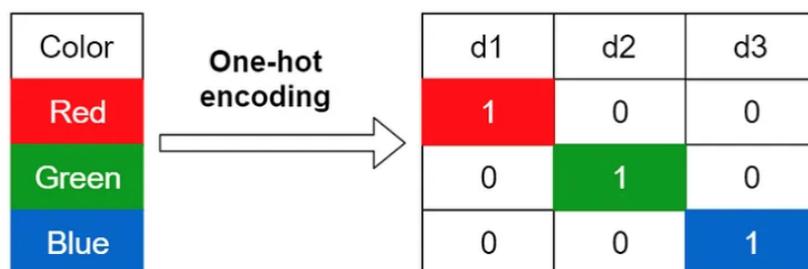
Figura 1 – Geração das word embeddings



Fonte: (Tripathi 2022).

A escolha de utilizar word embeddings em modelos baseados em Transformers, em detrimento de outros métodos, é fundamentada em diversas considerações desde a eficácia na representação semântica até a eficiência computacional. Por exemplo, na abordagem one-hot, o processo de codificação, cada termo presente em um vocabulário é expresso por de um vetor binário (0 e 1) cuja dimensão é equivalente ao tamanho total do vocabulário. A Figura 2 demonstra a maneira como as palavras são representas nessa abordagem.

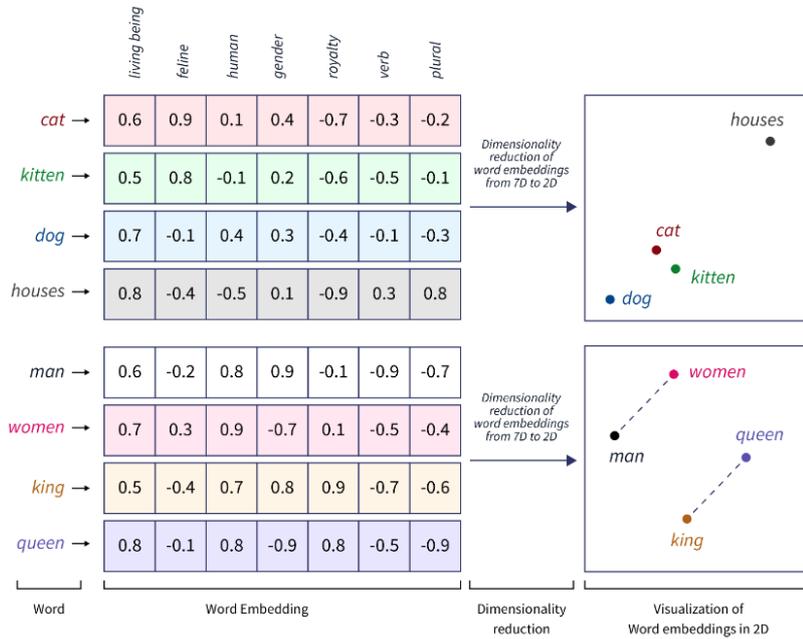
Figura 2 – Representação das palavras com codificação One-Hot.



Fonte: (Barreto 2023).

No exemplo fornecido, a abordagem de codificação one-hot resulta em vetores esparsos e ineficientes computacionalmente, incapazes de capturar nuances semânticas. Por outro lado, as word embeddings, ao serem representações vetoriais contínuas, conseguem apreender complexas relações semânticas, reduzindo a dimensionalidade de forma eficiente, favorecendo o aprendizado de contexto e a generalização, conforme ilustrado na Figura 3.

Figura 3 – Representação espacial das word embeddings



Fonte: (Malingan 2024).

### 2.3.1 Positional Encoder

Sabendo que os modelos utilizam vetores de embeddings para representar tokens ou palavras, a incorporação de uma codificação de posição é responsável pelo alinhamento adequado das posições das palavras com suas representações. Visando assegurar que cada posição tenha uma representação única, os autores (Vaswani et al. 2023) propõem o uso de funções senoidais e cossenoidais normalizadas entre  $[-1, 1]$  na geração de um vetor distinto para cada posição na sequência.

Cada vetor de codificação possui a mesma dimensão,  $d_{\text{model}}$ , que um vetor de embeddings, permitindo que ambos sejam somados. A variável  $k$  representa a posição, iniciando em 0 e indo até  $L - 1$ . Vale notar que o valor máximo que  $k$  pode assumir é  $d_{\text{model}}/2$ , dado que as equações alternam-se para cada elemento de embeddings. A variável  $n$  pode ser definida para qualquer valor, no entanto, os autores (Vaswani et al. 2023) recomendam até 1000. As Equações 1 e 2 descrevem o mecanismo de codificação posicional:

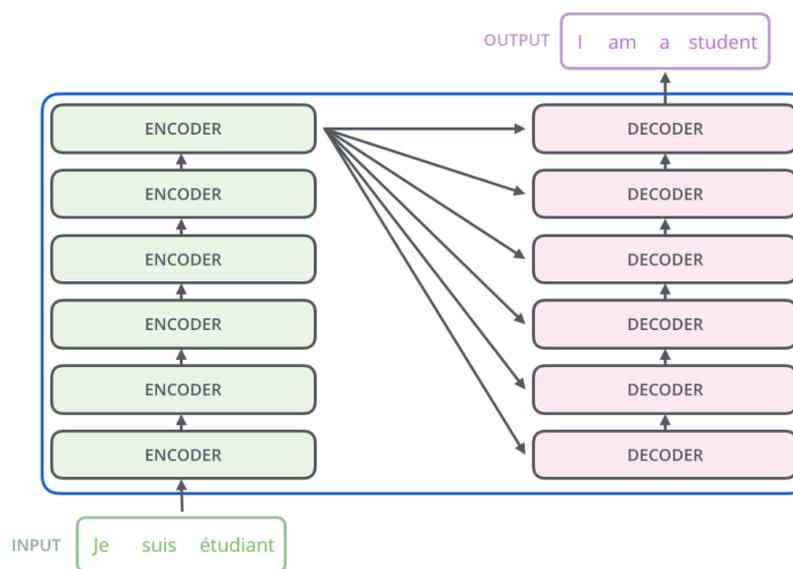
$$PE_{(pos,2i)} = \sin\left(\frac{pos}{10000(2i/d_{\text{model}})}\right) \quad (1)$$

$$PE_{(pos,2i+1)} = \cos\left(\frac{pos}{10000(2i/d_{\text{model}})}\right) \quad (2)$$

## 2.4 Transformers

Os Transformers, (Vaswani et al. 2023), representam uma arquitetura de rede neural inovadora destinada ao processamento de dados sequenciais, como texto ou séries temporais. Em contraste com abordagens recorrentes, essa arquitetura utiliza exclusivamente mecanismos de atenção para compreender as dependências globais entre a entrada e a saída. Seu desenvolvimento teve origem no âmbito da tradução de sequências de tokens, onde “tokens” refere-se a elementos individuais de uma sequência, como palavras em uma frase.

Figura 4 – Arquitetura Transformers Simplificada



Fonte: (Alammar 2022).

A Figura 4 ilustra a arquitetura simplificada dos Transformers, exemplificando sua aplicação na tradução de texto do francês para o inglês. O Encoder, responsável pela fase inicial, processa a sequência de palavras na língua de origem (francês), extraíndo representações de alta qualidade para cada palavra por meio do mecanismo de self-attention. Esse mecanismo permite a codificação de cada palavra considerando suas interações com todas as outras na sentença original, proporcionando uma compreensão profunda do contexto. A representação final gerada pelo Encoder condensa informações relevantes da sentença original.

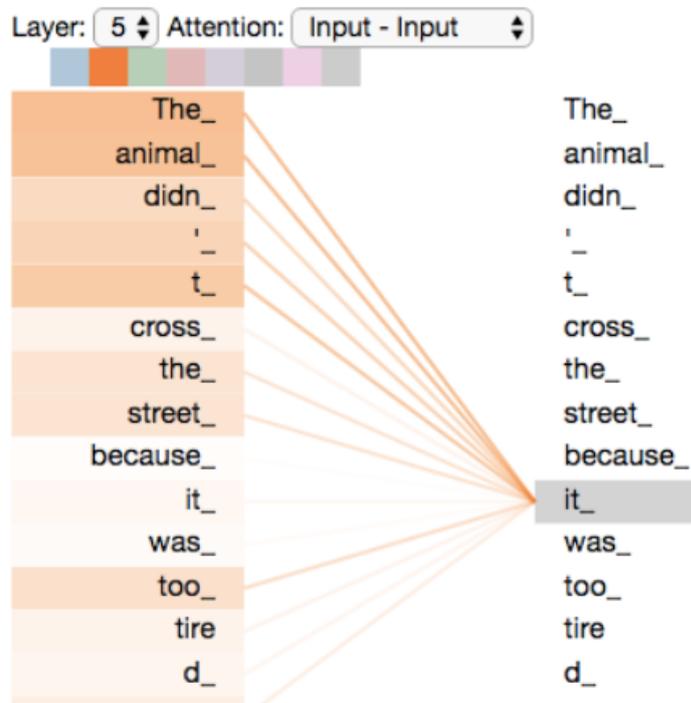
O Decoder, por sua vez, recebe a representação do Encoder e utiliza uma atenção mais complexa para gerar a tradução na língua alvo (inglês). Similar ao Encoder, o Decoder aplica o self-attention, mas adiciona a atenção ao contexto da sentença de origem. Isso possibilita ao modelo “olhar” para diferentes partes da sentença original ao gerar cada palavra da tradução. Cada palavra no Decoder é gerada condicionalmente com base nas informações acumuladas até o momento, proporcionando uma tradução coesa

e significativa. Essa abordagem demonstra o poder dos Transformers na compreensão e geração de sequências, tornando-os fundamentais em tarefas como a tradução automática.

### 2.4.1 Attention

Conforme previamente discutido, os Transformers empregam um mecanismo de atenção. Quando aplicado à tradução, o propósito é computar as relações individuais entre cada palavra em uma frase e todas as outras palavras da mesma sentença. Esse procedimento conduz a uma análise das interações contextuais entre as palavras. A Figura 5 traz uma representação visual do processo de cálculo da atenção para a palavra “it”, levando em consideração todos os demais tokens presentes na frase.

Figura 5 – Ilustração do mecanismo de self-attention



Fonte: (Alammar 2022).

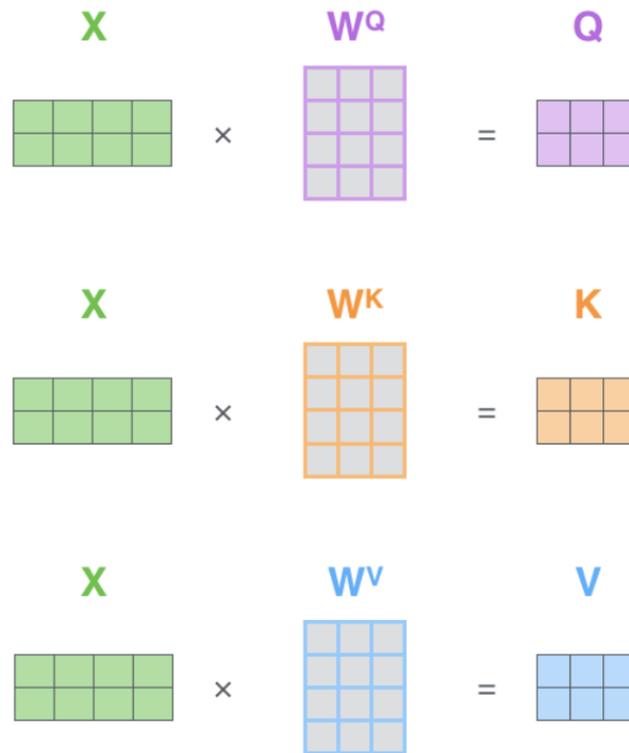
No âmbito dos Transformers, a função de atenção mapeia uma “query” ( $Q$ ) em relação a um conjunto de pares de chaves e valores ( $K$ - $V$ ). Essa função resulta em uma saída ponderada, na qual os pesos são determinados pela medida de relevância entre a “query” e as chaves associadas. A formalização matemática que encapsula essa operação é expressa pela seguinte equação:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) \cdot V \quad (3)$$

Na Figura 6, é apresentado o processo de geração dos vetores  $Q$ ,  $K$  e  $V$  por meio da multiplicação da matriz de embeddings com os vetores de peso. Esse passo é essencial para

a aplicação do mecanismo de atenção, onde a qualidade desses vetores influencia diretamente na capacidade do modelo em capturar as relações entre as palavras da sequência.

Figura 6 – Geração dos vetores Q e (K-V).



Fonte: (Alammar 2022).

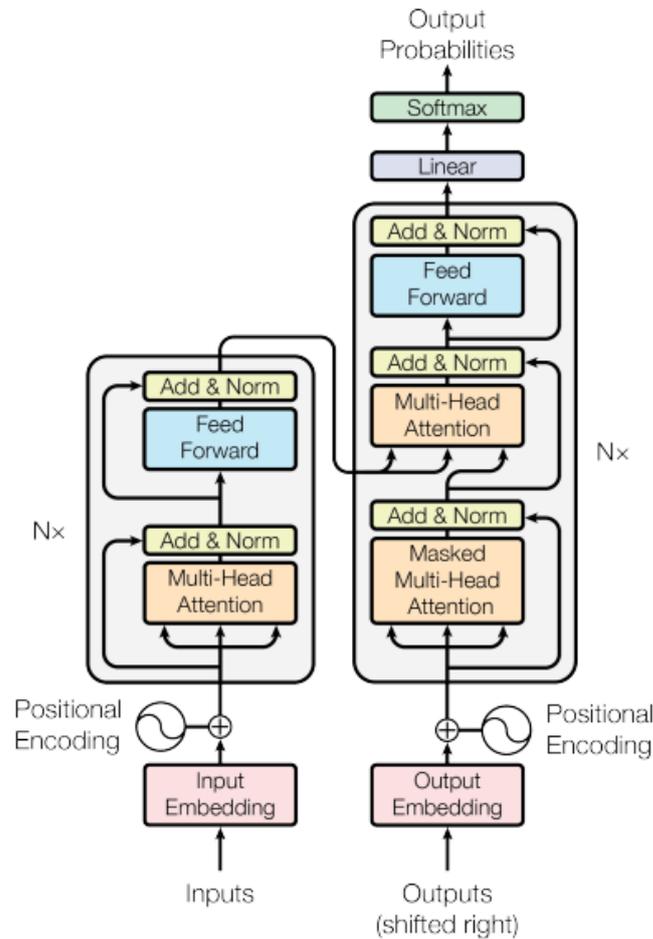
Por fim, calcula-se o resultado da camada de atenção usando a Equação (3), como ilustrado na Figura 7.

Figura 7 – Cálculo do mecanismo de atenção

$$\text{softmax} \left( \frac{\begin{matrix} Q \\ \times \\ K^T \end{matrix}}{\sqrt{d_k}} \right) \begin{matrix} V \\ \\ \end{matrix} = \begin{matrix} Z \\ \\ \end{matrix}$$

Fonte: (Alammar 2022).

Figura 8 – Arquitetura de Transformers em Alto Nível



Fonte: (Vaswani et al. 2023).

## 2.4.2 Encoder-Decoder

A arquitetura encoder-decoder de Transformers é uma estrutura de rede neural que usa os módulos encoder-decoder dos transformares para execução de várias tarefas de PLN. De maneira simples, o encoder processa a sequência de entrada e produz uma representação contínua da entrada, que então passada ao decoder para gerar a sequência de saída. A Figura 8 representa a arquitetura encoder-decoder em auto nível.

### 2.4.2.1 Encoder

No encoder, representado como a porção esquerda da Figura 8, os dados de entrada são convertidos em representações vetoriais por meio de uma camada de embeddings, que é inicializada com embeddings de palavras pré-treinadas. Posteriormente, para aprimorar a representação, os dados são enriquecidos por uma codificação posicional, utilizando funções senoidais conforme demonstrado na seção anterior.

A etapa subsequente, conhecida como multi-head-attention, emprega mecanismos de

self-attention para capturar dependências entre diferentes partes dos dados. Após a execução da self-attention, os vetores resultantes são agregados à conexão residual da camada de entrada e, em seguida, normalizados. Subsequentemente, a saída da self-attention é encaminhada através de uma rede de avanço ponto a ponto, composta por duas camadas lineares com uma função de ativação não linear interposta. Ao finalizar este processo, a saída é novamente submetida à normalização, com conexões residuais adicionadas da camada anterior, permitindo que o modelo assimile informações das camadas anteriores e do processo de avanço ponto a ponto. Este conjunto de operações proporciona ao encoder a capacidade de capturar relações complexas e contextuais nos dados de entrada.

#### 2.4.2.2 Decoder

No decoder, representado como a porção direita da Figura 8, a saída da etapa anterior é utilizada como entrada, iniciando o processo com uma camada de embeddings, semelhante ao encoder. Essa camada converte os tokens de saída em representações vetoriais e é inicializada com embeddings de palavras pré-treinadas.

Para assegurar que o decoder não acesse informações de tokens futuros na sequência de entrada durante a geração de saída, é empregado um mecanismo de masked multi-head attention. A etapa de atenção mascarada é sucedida por uma normalização, seguindo o mesmo padrão do processo no encoder. Além da atenção mascarada, o bloco do decoder inclui conexões residuais para facilitar o fluxo de gradientes e aprimorar o desempenho global. A saída da atenção mascarada é então processada por uma rede de avanço ponto a ponto, composta por duas camadas lineares com uma função de ativação não linear entre elas.

A saída da rede de avanço ponto a ponto passa por uma normalização adicional, com conexões residuais adicionadas da camada anterior. Esse design possibilita que o modelo incorpore informações das camadas anteriores e do processo de avanço ponto a ponto.

## 2.5 BERT

O Bidirectional Encoder Representations from Transformers (BERT), representa uma evolução significativa em relação aos modelos de representação textual anteriores, introduzindo uma abordagem bidirecional para a compreensão contextual de palavras em uma sequência. Em contraste com os modelos unidirecionais, o BERT processa todas as palavras em uma sentença simultaneamente, capturando relações complexas e dependências contextuais.

A arquitetura do BERT proposta em (Devlin et al. 2019) destaca-se pela atenção multi-head e camadas de autoatendimento, permitindo uma representação mais rica e contextual das palavras. A estrutura autorregressiva, presente em modelos sequenciais

tradicionais, é substituída pela abordagem bidirecional do BERT, habilitando o modelo a entender o contexto global de cada palavra em um texto.

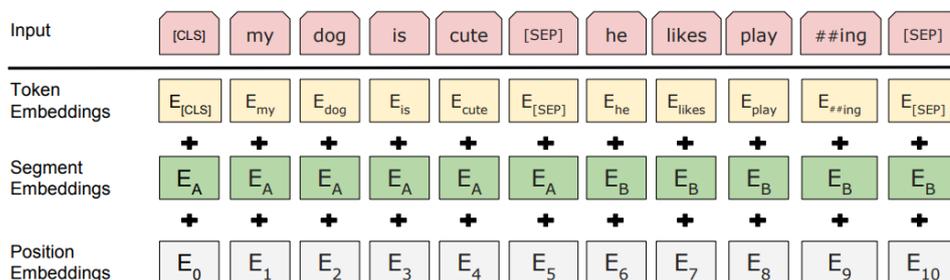
### 2.5.1 Arquitetura do Modelo

O BERT se apoia em uma estrutura proveniente do Transformers. Dado que o BERT está focado na geração de um modelo de representação da linguagem, ele requer exclusivamente a utilização do componente de codificação do Transformer. Nesse contexto, a entrada para o codificador consiste em uma sequência de tokens, os quais são primeiramente convertidos em vetores para subsequente processamento na rede neural. Entretanto, antes que o processamento propriamente dito ocorra, é essencial adicionar alguns metadados à entrada do modelo:

1. **Incorporações de Tokens:** Introdz-se um token [CLS] no início dos tokens de palavras da primeira frase, enquanto um token [SEP] é inserido ao final de cada frase. Essa estratégia visa proporcionar uma estrutura inicial e final distintiva à sequência de tokens.
2. **Incorporações de Seguintos:** Um marcador para distinção das sentenças é inserido, isso permite que o decodificador diferencie a sentença A da sentença B.
3. **Incorporações de Posições:** Um incorporação posicional é adicionada a cada token para indicar a sua posição na frase.

A Figura 9 ilustra o processo de entrada dos tokens ou palavras até a geração dos codificadores de posição para o modelo BERT.

Figura 9 – Representação das Entradas das Sentenças no BERT.



Fonte: (Devlin et al. 2019).

Na fase inicial, denominada entrada (input), é apresentada uma sequência de textos ao codificador como parte do processo. Posteriormente, cada token da sequência é convertido em uma representação numérica, expressa como um vetor contendo números reais (token embeddings). Após a obtenção desses embeddings para cada token na sequência, torna-se possível gerar uma representação abrangente para a frase como um todo, isso

alcançado mediante a combinação dos embeddings individuais associados aos tokens que compõem a sequência (segment embeddings). Por fim, os embeddings posicionais (positional embeddings) são adicionados aos embeddings de token para levar em conta a ordem e a posição das palavras na sequência.

### 2.5.2 Modelagem de Linguagem Mascarada

A ideia desta abordagem é aparentemente simples: aleatoriamente mascarar 15% das palavras na entrada substituindo-as por um token [MASK], submeter a sequência completa ao codificador baseado em atenção do BERT e, posteriormente, predizer apenas as palavras mascaradas, com base no contexto proporcionado pelas outras palavras não mascaradas na sequência. Contudo, surge um problema com essa abordagem de mascaramento, o modelo tenta prever apenas quando o token [MASK] está presente na entrada, enquanto almejamos que o modelo tente prever as palavras corretas, independentemente do token presente na entrada.

- 80% dos tokens são efetivamente substituídos pelo token [MASK].
- Em 10% das ocasiões, os tokens são substituídos por um token aleatório.
- Em 10% das ocasiões, os tokens permanecem inalterados.

Durante o treinamento, a função de perda do BERT considera apenas a predição dos tokens mascarados, ignorando a predição daqueles que não foram mascarados. Esse procedimento resulta em um modelo que converge significativamente mais lentamente do que os modelos de leitura da esquerda para a direita ou da direita para a esquerda.

## 2.6 RoBERTa

O A Robustly Optimized BERT Pretraining Approach (RoBERTa), introduzido pelos autores em (Liu et al. 2019) é considerado um variante do modelo BERT, pois compartilha da mesma arquitetura. Portanto, assim como o BERT, o RoBERTa é um modelo de linguagem baseado em Transformers que usa o mecanismo de attention para processar as sequenciais de entrada e gerar representações contextualizadas de uma palavra em uma frase. Todos os avanços trazidos pelo RoBERTa serão explicados nas próximas subseções.

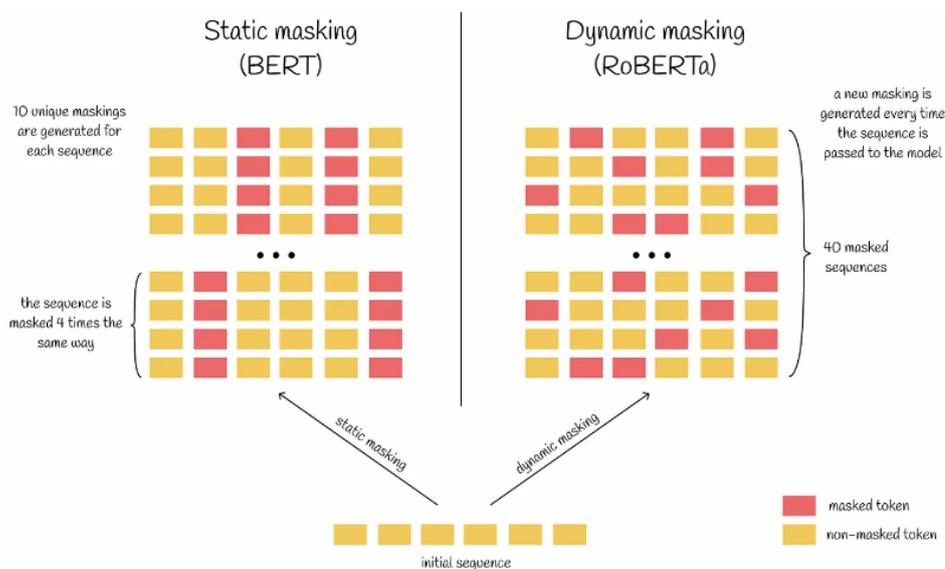
### 2.6.1 Dynamic Masking

Na arquitetura do BERT, durante a fase de pré-treinamento, o modelo realiza a modelagem de linguagem, tentando antecipar uma certa porcentagem de tokens mascarados.

Contudo, uma limitação na implementação original reside no fato de que os tokens escolhidos para mascaramento em uma sequência de texto podem ser os mesmos em diferentes lotes.

Para abordar essa questão, o conjunto de dados de treinamento é replicado 10 vezes, resultando em cada sequência sendo mascarada apenas de 10 maneiras distintas. Com o BERT sendo treinado por 40 épocas, cada sequência com o mesmo mascaramento é apresentada ao modelo quatro vezes. No entanto, pesquisas indicam que é ligeiramente mais eficaz utilizar um mascaramento dinâmico. Nesse método, o mascaramento é gerado de forma única a cada vez que uma sequência é submetida ao BERT. Globalmente, essa abordagem reduz a duplicação de dados durante o treinamento, proporcionando ao modelo a oportunidade de interagir com dados mais diversos e padrões de mascaramento. Para uma melhor compreensão, a Figura 10 ilustra a diferença entre o mascaramento estático e o mascaramento dinâmico.

Figura 10 – Mascaramento Estático x Dinâmico.



Fonte: (Horev 2018).

## 2.6.2 Otimizações

As principais otimizações introduzidas no RoBERTa de acordo com (Liu et al. 2019):

- ❑ **Duração do Pré-treinamento:** Originalmente o BERT varia entre 1.000.000 de steps enquanto o RoBERTa varia de 100.000 steps até 500.000 steps.
- ❑ **Conjunto de dados de Pré-treinamento:** Inclusão de dados adicionais, variando o tamanho do conjunto de dados de 16GB a 160GB.
- ❑ **Batch Size:** Originalmente o BERT tinha 256 seqüências, com o RoBERTa passa a ter 8000 seqüências.

- **Learning Rate e Otimizador:** O RoBERTa implementa um aquecimento da taxa de aprendizado nos primeiros 10.000 steps até um pico de  $1e-4$ , seguido de decaimento linear.

### 2.6.3 BERTimbau

O BERTimbau é uma adaptação do modelo BERT, especificamente pré-treinado para o idioma português brasileiro, criado pelos autores (Souza e Lotufo 2020). Este modelo alcançou resultados nível estado da arte em três tarefas de Processamento de Linguagem Natural (PLN) para o português como: Reconhecimento de Entidades Nomeadas (NER), Similaridade Textual de Sentenças (STS) e Inferência Textual.

A metodologia empregada para o desenvolvimento do BERTimbau replica de perto a arquitetura e os procedimentos de pré-treinamento do BERT, conforme detalhado na seção anterior. A principal distinção reside no corpus utilizado para a etapa de treinamento, que é composto exclusivamente por textos em português. Esta abordagem garante que o modelo esteja adequadamente adaptado às nuances e particularidades do idioma português brasileiro.

O modelo BERTimbau está disponível em duas versões: BERTimbau-Base (12 camadas, 768 dimensões ocultas, 12 cabeças de atenção e 110M de parâmetros) e o BERTimbau-Large (24 camadas, 1024 dimensões ocultas, 16 cabeças de atenção e 330M de parâmetros), ambos limitando o comprimento máximo da sentença em 512 tokens.



---

## Capítulo 3

# Trabalhos Relacionados

---

Neste capítulo, serão discutidas duas estratégias que utilizaram os modelos BERT e suas variantes para a tarefa de classificação de emoções. A primeira abordagem refere-se à aplicação específica do modelo BERTimbau na classificação de emoções, enquanto a segunda explora a avaliação de diversos modelos baseados na arquitetura de transformadores, também direcionados à classificação de emoções.

### 3.1 Utilizando BERTimbau para a Classificação de Emoções em Português

O estudo conduzido pelos pesquisadores (Hammes e Freitas 2021) envolve o processo de fine-tuning do modelo BERTimbau, adaptando-o para a tarefa específica de classificação de emoções. Os autores utilizaram o conjunto de dados GoEmotions, que foi traduzido para o português com o auxílio da biblioteca itranslate. Este conjunto de dados é composto por 54.263 sentenças, as quais foram manualmente anotadas em 28 classes, incluindo 27 emoções distintas e uma classe para sentenças neutras.

Devido ao desbalanceamento observado nas classes, os autores optaram por empregar uma técnica conhecida como Class Balanced Loss (CB) como uma solução para lidar com esse problema. Os resultados obtidos indicaram melhorias significativas ao utilizar o modelo BERTimbau, nas versões base e large, em comparação com o desempenho do modelo BERT-base, originalmente treinado para o conjunto de dados em inglês como pode ser visto na Figura 3.1.

Os resultados evidenciam que a implementação do modelo BERTimbau, em suas duas versões, aliada à técnica de balanceamento CB, resultam em melhorias significativas nos

Figura 11 – Comparativo entre os modelos para inglês e português.

Emoções	BERT-base EN [Devlin et al. 2018]			BERTimbau-base PT [Souza et al. 2020] Balanceado			BERTimbau-large PT [Souza et al. 2020] Balanceado		
	P	S	F1	P	S	F1	P	S	F1
	admiração	0,53	0,83	0,65	0,58	0,75	0,66	0,60	0,75
diversão	0,70	0,94	0,80	0,76	0,89	0,82	0,75	0,91	0,82
raiva	0,36	0,66	0,47	0,37	0,46	0,41	0,40	0,49	0,44
aborrecimento	0,24	0,63	0,34	0,37	0,33	0,35	0,35	0,38	0,36
aprovação	0,26	0,57	0,36	0,40	0,40	0,40	0,42	0,40	0,41
zelo	0,30	0,56	0,39	0,34	0,44	0,39	0,34	0,47	0,39
confusão	0,24	0,76	0,37	0,33	0,56	0,41	0,35	0,58	0,44
curiosidade	0,40	0,84	0,54	0,44	0,77	0,56	0,45	0,80	0,57
desejo	0,43	0,59	0,49	0,53	0,55	0,54	0,60	0,58	0,59
decepção	0,19	0,52	0,28	0,28	0,20	0,23	0,28	0,25	0,26
desaprovação	0,29	0,61	0,39	0,37	0,43	0,40	0,39	0,45	0,42
nojo	0,34	0,66	0,45	0,46	0,42	0,44	0,46	0,47	0,46
constrangimento	0,39	0,49	0,43	0,39	0,38	0,38	0,39	0,35	0,37
entusiasmo	0,26	0,52	0,34	0,36	0,43	0,39	0,35	0,50	0,41
medo	0,46	0,85	0,60	0,54	0,73	0,62	0,57	0,77	0,65
gratidão	0,79	0,95	0,86	0,88	0,91	0,90	0,88	0,92	0,90
luto	0,00	0,00	<b>0,00</b>	0,13	0,67	<b>0,22</b>	0,17	0,50	0,25
alegria	0,39	0,73	0,51	0,51	0,57	0,54	0,51	0,60	0,55
amor	0,68	0,92	0,78	0,72	0,87	0,79	0,70	0,85	0,77
nervosismo	0,28	0,48	0,35	0,29	0,48	0,36	0,24	0,43	0,31
neutro	0,56	0,84	0,68	0,64	0,70	0,67	0,64	0,70	0,67
otimismo	0,41	0,69	0,51	0,52	0,56	0,54	0,53	0,53	0,53
orgulho	0,67	0,25	0,36	0,39	0,44	0,41	0,36	0,50	0,42
percepção	0,16	0,29	0,21	0,37	0,12	0,18	0,34	0,21	0,26
alívio	0,50	0,09	0,15	0,14	0,27	0,18	0,18	0,55	0,27
remorso	0,53	0,88	0,66	0,51	0,88	0,64	0,54	0,88	0,67
tristeza	0,38	0,71	0,49	0,47	0,54	0,50	0,48	0,57	0,52
surpresa	0,40	0,66	0,50	0,49	0,60	0,54	0,46	0,58	0,51
<b>média macro</b>	<b>0,40</b>	<b>0,63</b>	<b>0,46</b>	<b>0,45</b>	<b>0,55</b>	<b>0,48</b>	<b>0,45</b>	<b>0,57</b>	<b>0,50</b>

Fonte: (Hammes e Freitas 2021).

resultados obtidos. Contudo, os pesquisadores sugerem que o desempenho poderia ter sido ainda mais aprimorado caso houvesse disponibilidade de ferramentas de tradução automática mais eficazes, pois foi encontrado incoerências nas traduções.

## 3.2 Fine-grained Sentiment Classification using BERT

No estudo conduzido por (Shakya e Shrestha 2019), foi executado o processo de fine-tuning nos modelos BERT nas versões base e large, com o propósito de realizar a classificação de emoções nos conjuntos de dados SST-2 e SST-5. O Stanford Sentiment Treebank (SST) destaca-se como um dos conjuntos de dados públicos mais amplamente reconhecidos e utilizados na tarefa de classificação de emoções. Este repositório abriga 11.855 avaliações cinematográficas, onde cada frase é submetida à análise do analisador de constituintes de Stanford, resultando em uma estrutura de árvore, onde a frase completa é designada como o nó raiz ("root") e as palavras individuais são representadas como nós folha ("leaf"). Adicionalmente, cada nó dessa estrutura é meticulosamente rotulado por no mínimo três

avaliadores humanos. Esses rótulos abrangem cinco categorias sentimentais distintas, variando desde "muito negativo" até "muito positivo". Durante a fase de pré-processamento dos dados, foram aplicadas técnicas como remoção de dígitos, pontuações e acentos, além da conversão das palavras para minúsculas. Adicionalmente, empregou-se o tokenizador WordPiece para segmentação das palavras em prefixos, raízes e sufixos, proporcionando uma abordagem mais refinada ao tratamento de palavras não previamente vistas.

No processo de treinamento, os autores realizaram comparações entre diversos algoritmos, tais como CNN, BiLSTM, LSTM, RNTN, RNN, Paragraph Vectors e Avg word vectors. Os resultados, apresentados na Figura 3.2, indicam um desempenho superior para os modelos BERT em ambas as versões, evidenciando a eficácia desse modelo na tarefa de classificação de emoções nos conjuntos de dados SST-2 e SST-5.

Figura 12 – Comparativo entre os modelos para inglês e português.

Model	SST-2		SST-5	
	All	Root	All	Root
Avg word vectors [9]	85.1	80.1	73.3	32.7
RNN [8]	86.1	82.4	79.0	43.2
RNTN [9]	87.6	85.4	80.7	45.7
Paragraph vectors [2]	–	87.8	–	48.7
LSTM [10]	–	84.9	–	46.4
BiLSTM [10]	–	87.5	–	49.1
CNN [11]	–	87.2	–	48.0
<b>BERT<sub>BASE</sub></b>	<b>94.0</b>	<b>91.2</b>	<b>83.9</b>	<b>53.2</b>
<b>BERT<sub>LARGE</sub></b>	<b>94.7</b>	<b>93.1</b>	<b>84.2</b>	<b>55.5</b>

Fonte: (Shakya e Shrestha 2019).

Os resultados obtidos nesta pesquisa destacam a superioridade dos modelos BERT nas versões base e large em comparação com outros modelos avaliados. Essa superioridade é evidenciada tanto na classificação das sentenças completas (abordagem 'root') quanto na análise mais granular de cada nó individual na árvore sintática gerada para as sentenças (abordagem 'all').

### 3.3 Comparando os Trabalhos

Ambos os estudos investigam a classificação de emoções mediante a aplicação de técnicas de fine-tuning em modelos BERT ou suas variantes. No primeiro trabalho, o modelo BERTimbau é utilizado no conjunto de dados GoEmotion, adaptado para a língua portuguesa. Os resultados indicam melhorias significativas na tarefa de classificação de emoções, comparadas ao BERT-base aplicado ao conjunto GoEmotions em inglês. No segundo trabalho, os modelos BERT base e large são empregados na classificação de emoções nos

conjuntos de dados SST-2 e SST-5. Os resultados obtidos destacam a superioridade desses modelos em relação a outros algoritmos avaliados .

---

# Capítulo 4

## Desenvolvimento

---

Neste capítulo, será explicado o processo de desenvolvimento planejado para condução deste estudo. A exposição inclui a apresentação das ferramentas empregadas na construção do modelo e na etapa de pré-processamento dos dados. Em seguida, é realizada uma exposição detalhada sobre o conjunto de dados utilizado neste trabalho. Por fim, são discutidos os detalhes da utilização de cada modelo na fase de treinamento.

### 4.1 Ferramentas

Essa seção apresenta as ferramentas utilizadas no desenvolvimento do projeto.

#### 4.1.1 Tratamento de Dados

As ferramentas empregadas na fase de tratamento de dados consistiram em:

1. **Normalizador Enelvo** (Bertaglia e Nunes 2016): Ferramenta construída em Python usada para normalização de textos em português.
2. **Regex**: O módulo nativo do Python que oferece suporte a operações com expressões regulares, foi empregado para manipular strings de maneira mais complexa. Através desse módulo, tornou-se viável remover caracteres indesejados e pontuações das sentenças, contribuindo para a limpeza e padronização dos dados textuais.

### 4.1.2 Criação de Modelos

O framework utilizado para construção dos modelos foi o PyTorch. O PyTorch é uma biblioteca de código aberto para aprendizado de máquina e inteligência artificial. Desenvolvido principalmente pelo Facebook's AI Research lab (FAIR), o PyTorch fornece uma estrutura flexível e dinâmica para a construção de modelos de aprendizado profundo.

Outro framework utilizado foi a biblioteca Transformers, desenvolvida pela Hugging Face, que é uma biblioteca em Python para trabalhar com modelos de linguagem pré-treinados no campo de Processamento de Linguagem Natural (PLN) e aprendizado profundo. Essa biblioteca é conhecida por sua extensa coleção de modelos pré-treinados, incluindo modelos como BERT, GPT (Generative Pre-trained Transformer), RoBERTa, BERTimbau entre outros. Para mais detalhes, todos os códigos usados neste projeto para à criação e treinamento dos modelos estão disponíveis no Github<sup>1</sup>.

## 4.2 Base de Dados

O B2W-Reviews, constitui um corpus aberto de avaliações de produtos, compreendendo mais de 130 mil avaliações de clientes de ecommerce, coletadas do site Americanas.com entre janeiro e maio de 2018. Esse corpus proporciona informações detalhadas sobre o perfil do consumidor, incluindo gênero, idade e localização geográfica. Além disso, apresenta duas formas distintas de avaliação:

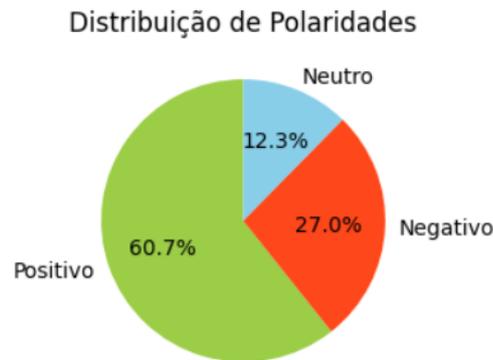
1. A tradicional escala de classificações de 5 pontos, representada por estrelas na maioria dos ecommerce.
2. Um rótulo “recomendar a um amigo”, que é uma pergunta de “sim ou não” que reflete a disposição do cliente em recomendar o produto a outra pessoa.

A escolha de utilizar a escala de 5 estrelas como característica-alvo, foi motivada pelo interesse em examinar o processo de aprendizado dos modelos diante de distintas polaridades em um paradigma de três classes (positivo, negativo e neutro). Essa escolha busca não apenas avaliar a habilidade do modelo em identificar a classe neutra, mas, de forma concomitante, analisar a sua aptidão para discernir as diferentes classes. Dessa maneira, a polaridade positiva se correlaciona com avaliações de 4 e 5 estrelas, ao passo que a polaridade negativa está associada a avaliações de 1 e 2 estrelas; as avaliações de 3 estrelas, por sua vez, são consideradas como instâncias de polaridade neutra.

A Figura 13 representa a distribuição de classes presentes no corpus B2W-Reviews.

<sup>1</sup> <<https://github.com/g-barbosa11/sentiment-classifiers>>

Figura 13 – Distribuição de Polaridades



Fonte: (Autoria Própria)

## 4.3 Pré-Processamento

Esta seção descreve o pré-processamento aplicado no corpus B2W-Reviews preparando-o para etapa de treinamento. O processamento foi aplicado na seguinte ordem:

1. **Normalização com o Enelvo:** Com o normalizador foi possível fazer a correção de sentenças discrepantes, remoção de emojis e hashtags.
2. **Remoção de Caracteres Indesejados por meio da Regex:** Remoção de números ou símbolos que não contribuem para a compreensão do texto.
3. **Remoção de Pontuação por meio da Regex:** Remoção de pontuações e vírgulas.
4. **Geração de Word Embeddings:** Representação das palavras como vetores numéricos a partir de embeddings pré-treinado contidos nos próprios modelos utilizados neste trabalho.

## 4.4 Modelos Base

Foram propostas três soluções para a classificação de sentimentos, utilizando o fine-tuning dos modelos BERT-base<sup>2</sup>, RoBERTa-base<sup>3</sup> e BERTimbau-base<sup>4</sup>.

### 4.4.1 BERT e BERTimbau

Na etapa de configuração de ambos os modelos BERT e BERTimbau, foram inicialmente gerados os vetores de *attention mask*. Esses vetores indicam a validade dos

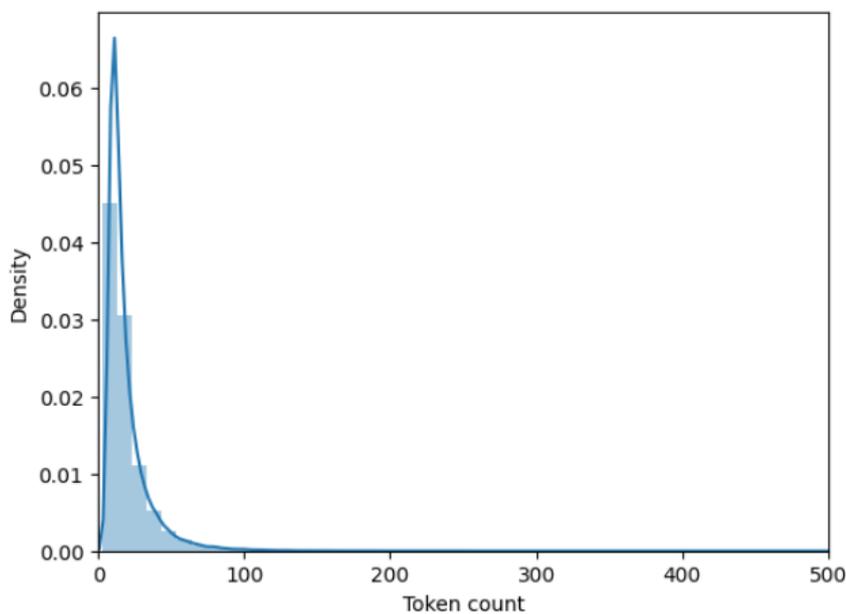
<sup>2</sup> <<https://huggingface.co/bert-base-multilingual-cased>>

<sup>3</sup> <<https://huggingface.co/FacebookAI/xlm-roberta-base>>

<sup>4</sup> <<https://huggingface.co/neuralmind/bert-base-portuguese-cased>>

elementos na sequência, destacando a atenção para os tokens relevantes durante o processamento. Além disso, foram introduzidos tokens especiais, como [SEP] para marcar o final de uma frase, [CLS] colocado no início da frase, que é crucial em problemas de classificação, e [PAD], que consiste em tokens de valor nulo. A inserção de [PAD] tem como objetivo assegurar que todas as sentenças tenham o mesmo tamanho, uma vez que o BERT opera com entradas de tamanho fixo. O parâmetro *max\_length* foi definido como 200, através da análise da distribuição da densidade de tokens no corpus B2W-Reviews, como ilustrado na Figura 4.4.1

Figura 14 – Distribuição da densidade de tokens no B2W-Reviews



Fonte: (Autoria Própria).

O classificador adotado consiste em uma camada de *dropout*, utilizada como mecanismo de regularização para mitigar o *overfitting* durante o treinamento, e uma camada linear de saída. Esta última é responsável pela projeção da representação resultante do BERT para o espaço de classes, facilitando a tarefa de classificação.

#### 4.4.2 RoBERTa

Na fase de configuração do modelo RoBERTa, foi adotado uma abordagem semelhante à utilizada nos modelos BERT e BERTimbau, com a diferença da introdução de uma camada adicional de ativação ReLU após a camada linear de projeção *pre\_classifier*, acrescentando não linearidade à representação. Subsequentemente, aplicamos uma camada de *dropout* como mecanismo de regularização durante o treinamento. A última camada do classificador é a camada linear de saída *classifier*, responsável por projetar a representação resultante do RoBERTa para o espaço de classes.

## 4.5 Treino

Esta seção descreve como foi feito o treinamento dos modelos propostos.

### 4.5.1 Dados de Entrada

Considerando o desbalanceamento de classes no conjunto de dados B2W-Reviews, evidenciado pela representação inferior das classes neutras conforme ilustrado na Figura 13, optou-se por selecionar quantidades equivalentes de instâncias para as classes positivas e negativas, correspondendo à quantidade de instâncias neutras presentes no corpus. Essa estratégia visa criar uma condição de treinamento na qual as classes apresentem um equilíbrio, proporcionando uma representação mais justa no processo de aprendizado dos modelos. Por fim o conjunto de dados foi dividido em em 80% para treinamento, 10% para teste e 10% para validação.

### 4.5.2 Loss Function

A *loss function* utilizada foi a *Categorical cross-entropy*.

### 4.5.3 Hiperparâmetros

Os hiperparâmetros foram selecionados de forma manual, tomando como referência o estudo apresentado por (Hammes e Freitas 2021). Neste trabalho, os autores empregaram o modelo BERTimbau para realizar a classificação de emoções utilizando o conjunto de dados GoEmotions, que foi traduzido para o idioma português. Os valores específicos atribuídos aos hiperparâmetros para todos os modelos analisados estão documentados na Tabela 1. Contudo, é relevante destacar que, no caso do modelo RoBERTa, optou-se por não utilizar os parâmetros relacionados ao learning rate scheduler e à proporção de warmup.

Tabela 1 – Valores dos Hiperparâmetros

Hiperparâmetro	Valor
<i>epochs</i>	4
<i>batch_size</i>	16
<i>max_length</i>	200
<i>optimizer</i>	AdamW
<i>learning rate</i>	2e-5
<i>learning rate scheduler</i>	linear with warmup
<i>warmup proportion</i>	0,2

Fonte: (Autoria Própria)

## 4.6 Métricas de Avaliação

Esta seção descreve as métricas utilizadas para a avaliação das abordagens propostas na tarefa de classificação de sentimentos. Ao metrificar os resultados pelos valores previstos para a polaridade, essas métricas indiretamente avaliam a qualidade das abordagens implementadas.

Na descrição das métricas a seguir, as siglas representam:

□ **VP**: Verdadeiro Positivo

□ **VN**: Verdadeiro Negativo.

□ **FP**: Falso Positivo.

□ **FN**: Falso Negativo.

**Acurácia (A)**: A acurácia é a razão entre as classificações corretas e o total de predições (soma das classificações corretas e incorretas).

$$\mathbf{A} = \frac{\mathbf{VP} + \mathbf{VN}}{\mathbf{VP} + \mathbf{VN} + \mathbf{FP} + \mathbf{FN}} \quad (4)$$

**Precisão (P)**: A precisão é a razão entre as polaridades classificadas corretamente e o total de polaridades preditas corretas e incorretas.

$$\mathbf{P} = \frac{\mathbf{VP}}{\mathbf{VP} + \mathbf{FP}} \quad (5)$$

**Revocação (R)**: A revocação é a razão entre as polaridades que foram corretamente classificadas e o total geral de polaridades da amostra.

$$\mathbf{R} = \frac{\mathbf{VP}}{\mathbf{VP} + \mathbf{FN}} \quad (6)$$

**Medida F1 (F)**: A medida F1 busca o equilíbrio entre precisão e revocação, utilizando uma média harmônica entre elas.

$$\mathbf{F} = 2 \times \frac{\mathbf{P} \times \mathbf{R}}{\mathbf{P} + \mathbf{R}} \quad (7)$$

Além dessas métricas, incorporou-se a matriz de confusão, uma ferramenta valiosa para uma análise mais detalhada do desempenho dos modelos de classificação. A matriz de confusão organiza as previsões em uma tabela, destacando os verdadeiros positivos (VP), verdadeiros negativos (VN), falsos positivos (FP) e falsos negativos (FN), como ilustrado na Figura 4.6.

Figura 15 – Comparativo entre os modelos para inglês e português.

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Fonte: (Narkhede 2018).



---

## Capítulo 5

# Experimentos e Resultados

---

Neste capítulo, são apresentados os resultados obtidos a partir do fine-tuning dos modelos BERT, BERTimbau e RoBERTa, nas suas versões base, utilizando o corpus B2W-Reviews. A performance dos modelos foi medida com as métricas de avaliação mencionadas no capítulo anterior.

### 5.1 Métricas dos Modelos

A Tabela 2 apresenta os resultados obtidos por cada um dos modelos testados. De maneira geral, os modelos demonstraram desempenho satisfatório na classificação das três categorias de classes. Contudo, as análises revelam que o BERTimbau-base se destacou, superando tanto o BERT-base quanto o RoBERTa-base em todas as métricas avaliadas.

Tabela 2 – Métricas de Classificação para Cada Modelo

Classes	BERT-base			RoBERTa-base			BERTimbau-base		
	P	R	F1	P	R	F1	P	R	F1
Positivo	0,78	0,80	0,79	0,80	0,79	0,80	0,81	0,82	0,81
Negativo	0,81	0,87	0,84	0,83	0,89	0,86	0,86	0,88	0,87
Neutro	0,68	0,62	0,65	0,68	0,64	0,66	0,71	0,67	0,69

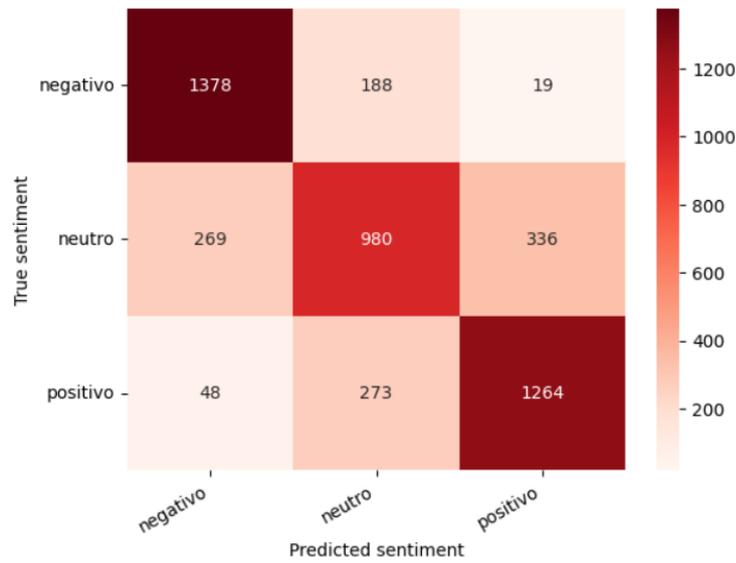
Fonte: (Autoria Própria)

### 5.2 Matrizes de Confusão

A análise das matrizes de confusão para cada um dos modelos proporciona um entendimento sobre como a classe "neutra" está sendo erroneamente classificada. Observa-se

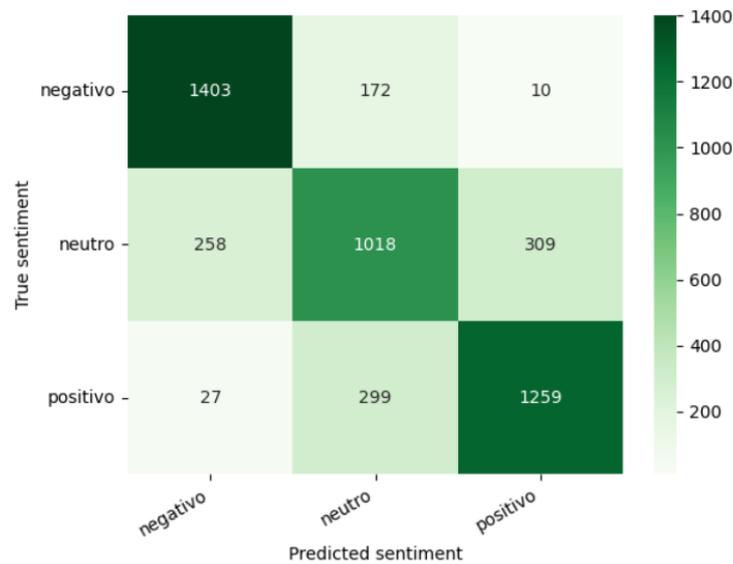
que os modelos têm a tendência de categorizar as mensagens neutras de forma mais positiva do que negativa, sugerindo uma inclinação para interpretar mensagens neutras como possuindo uma carga emocional mais positiva.

Figura 16 – Matriz de Confusão do Modelo BERT



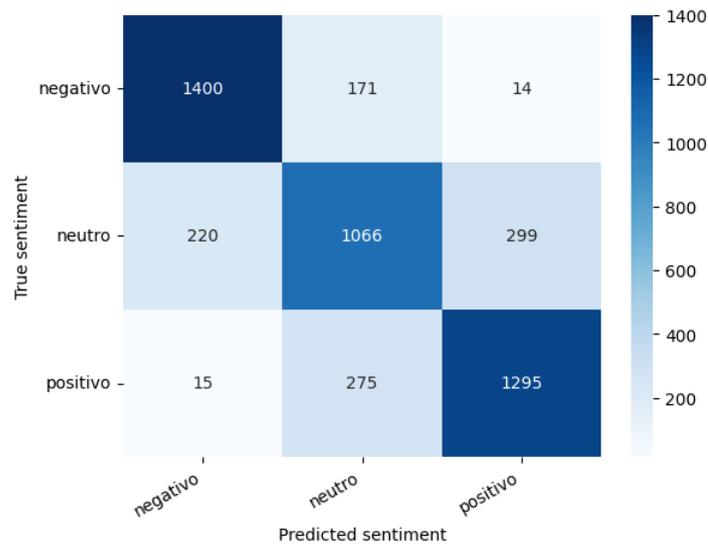
Fonte: (Autoria Própria)

Figura 17 – Matriz de Confusão do Modelo RoBERTa



Fonte: (Autoria Própria)

Figura 18 – Matriz de Confusão do Modelo BERTimbau



Fonte: (Autoria Própria)

Essa inclinação pode ser justificada pela predominância da classe positivas em relação às demais classes, conforme ilustrado na 13. Nos casos em que o modelo classifica uma sentença como negativa, embora esteja rotulada como positiva no conjunto de dados, observa-se uma discrepância entre a previsão do modelo e os rótulos originais das avaliações. Da mesma forma, quando o modelo classifica uma sentença como positiva, mas a classe real é negativa, a discrepância é evidenciada novamente, sugerindo uma inconsistência nas anotações do conjunto de dados. Na próxima seção, será apresentado exemplos que ilustram essa inconsistência encontrada em ambas previsões.

### 5.3 Exemplos de Classificação

Com base nos resultados apresentados pelas matrizes de confusão, uma análise das previsões do modelo e da qualidade do corpus B2W-Reviews pode ser realizada. Para os valores tabelados, foram usadas as previsões do BERTimbau-base como referência de valores previstos.

A Tabela 3 apresenta as instâncias em que o modelo classificou erroneamente as revisões como negativas, quando, na verdade, no corpus B2W-Reviews, elas estavam rotuladas como positivas (ou vice-versa), ocorreu uma discrepância nos rótulos de ambas as instâncias. Este cenário revela uma discrepância nos rótulos dos dados, indicando que o modelo, em determinados casos, classifica corretamente, mas em outros, incorre na classificação.

Ao analisar os dados na Tabela 4, observa-se que, para uma pontuação de 4 estrelas, as frases incorporam elementos tanto positivos quanto negativos, mas com uma inclinação predominantemente favorável ao positivo. Analogamente, ao lidar com pontuações de 2 estrelas, percebe-se que as frases tendem a ser mais negativas do que positivas. Essa

Tabela 3 – Exemplos Errôneos na Rotulagem do Corpus B2W-Reviews

Reviews	Rating	Previsão
não compraria nesta loja de novo; passamos a noite de natal sem o presente	5	Negativo
comprei com o benefício de ganhar reais em cápsulas e até o momento não consigo falar na americanas; péssimo atendimento pós-venda; não consigo cadastrar o benefício no site da dolce gusto; eles falam que a promoção era para se cadastrar até dezembro, mas a cafeteira chegou em janeiro e não veio nada de como ganhar os reais em cápsulas; estou decepcionado	5	Negativo
adorei o livro; é muito bom para quem começa a estudar anatomia, basicamente do zero; recomendo muito	1	Positivo
ótimo custo-benefício; aparelho moderno e está atendendo às minhas expectativas	1	Positivo

Fonte: (Autoria Própria)

dualidade apresenta um desafio ao modelo, dificultando sua capacidade de discernir com precisão a polaridade dessas avaliações, resultando em classificações inadequadas dessas expressões como neutras.

Tabela 4 – Exemplos de Erros Cometidos pelo Modelo ao Prever a Classe como Neutra

<b>Reviews</b>	<b>Rating</b>	<b>Previsão</b>
gostei muito do produto só acho que poderia melhor no cinto de segurança que não é confortável e não prende muito bem o bebê, mas muito bom o produto recomendo sim o produto	4	<b>Neutro</b>
o aparelho chegou dentro do prazo é compacto tem uma potência de som bom para ambientes domésticos; único ponto negativo é que descarrega com facilidade mesmo estando conectado à tomada	4	<b>Neutro</b>
gostei do produto, chegou um pouco atrasado; creio que por causa dos correios, mas o som é bom; a caixa de som é bem bonitinha, adorei	4	<b>Neutro</b>
o produto cumpre o que promete; é leve e fácil de usar, e realmente deixa os fios com aspecto de escovados; alisa cabelos ondulados, mas não recomendo para cabelos cacheados; poderia ser mais rápido na secagem dos cabelos	4	<b>Neutro</b>
produto bom, porém após comprar, comparei com a da bandeirantes, que usa peças e tecnologias superiores; não possui uma fácil regulagem de freios com a bandeirantes e não usa materiais robustos como a bandeirantes; pena eu ter visto isso só após a compra, mas não deixa de ser uma boa bicicleta	2	<b>Neutro</b>
adquiri o produto, mas não traz o que promete; as caixas acústicas até que transmitem um som de qualidade, porém a subwoofer deixa a desejar; tenho o modelo fama n, que apesar de menor, potência, tem muito mais qualidade de som	2	<b>Neutro</b>
boa qualidade, mas veio com apenas um soquete; portanto, não atende às minhas necessidades; pensei que fosse de lâmpadas e que viessem também as lâmpadas, mas não foi o caso	2	<b>Neutro</b>
layout lindo, porém com qualidade de som limitada; a parte de cima, similar a caixa da bateria convencional, emite um som sem vida; a parte de baixo, correspondente a um bumbo, o som já é calorzinho; pelo preço que é baixo, até vale a pena, mas para quem quer um bom som, talvez não	2	<b>Neutro</b>

Fonte: (Autoria Própria)



---

## Conclusão

---

Neste trabalho, foi realizado o *fine-tuning* de três modelos diferentes de redes neurais, baseados na arquitetura de transformers, adaptando-os especificamente para a classificação de sentimentos em opiniões de consumidores no contexto dos e-commerce. Inicialmente, procedeu-se com a coleta e pré-processamento do corpus B2W-Reviews, resultando na formação de um conjunto de dados relevante e adequado para esta pesquisa. Em seguida, foram empregados os três modelos distintos de redes neurais para realizar a classificação de sentimentos.

A avaliação dos resultados desses modelos foram interpretadas com uso de métricas de precisão, recall e f1-score, juntamente como o entendimento das matrizes de confusão. Neste contexto, o modelo BERTimbau mostrou sua superioridade em prever as três principais polaridades (positiva, negativa e neutra) destacando sua proeminência como um modelo de linguagem poderoso que pode ser ajustado para diversas tarefas de Processamento de Linguagem Natural (PLN) específica para o idioma português.

A constante evolução das arquiteturas transformers e a adaptação a domínios específicos, como o e-commerce, prometem contribuir para avanços notáveis na compreensão e interpretação do conteúdo emocional presentes na linguagem natural. Nesse sentido, a busca pela excelência na análise de sentimentos permanece um campo frutífero e repleto de oportunidades para inovação e aprimoramento contínuo.

Para pesquisas futuras, recomenda-se explorar a combinação de outros conjuntos de dados no domínio do e-commerce, ao corpus com o intuito de ampliar a representatividade das classes previstas. Além disso, sugere-se a conversão do problema multiclasse em uma série de problemas binários como uma estratégia promissora. Por exemplo, ao prever entre várias classes (A, B, C), é viável treinar três classificadores binários distintos: um para distinguir A de (B, C), outro para distinguir B de (A, C), e um terceiro para distinguir C de (A, B). A combinação das saídas desses classificadores pode resultar em previsões mais precisas. Adicionalmente, recomenda-se a implementação de ensemble dos modelos, combinando diversos classificadores a fim de alcançar previsões mais robustas

das classes. Essas recomendações têm o potencial de contribuir significativamente para o aprimoramento dos modelos de previsão em tarefas multiclasse no âmbito do e-commerce, incrementando tanto a precisão quanto a capacidade de generalização dos mesmos.

---

# Referências

---

ALAMMAR, J. **The Illustrated Transformer**. 2022. Acesso em 28 de janeiro de 2024. Disponível em: <<https://jalamar.github.io/illustrated-transformer/>>.

BARRETO, J. D. One-hot encoding in pharo: A powerful data preprocessing technique. **LatinXinAI**, July 2023. Disponível em: <<https://medium.com/latinxinai/one-hot-encoding-in-pharo-a-powerful-data-preprocessing-technique-516e481ab604>>.

BERTAGLIA, T. F. C.; NUNES, M. d. G. V. Exploring word embeddings for unsupervised textual user-generated content normalization. In: **Proceedings of the 2nd Workshop on Noisy User-generated Text (WNUT)**. [S.l.: s.n.], 2016. p. 112–120.

DEVLIN, J. et al. BERT: Pre-training of deep bidirectional transformers for language understanding. In: BURSTEIN, J.; DORAN, C.; SOLORIO, T. (Ed.). **Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)**. Minneapolis, Minnesota: Association for Computational Linguistics, 2019. p. 4171–4186. Disponível em: <<https://aclanthology.org/N19-1423>>.

HAMMES, L.; FREITAS, L. Utilizando bertimbau para a classificação de emoções em português. In: **Anais do XIII Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana**. Porto Alegre, RS, Brasil: SBC, 2021. p. 56–63. ISSN 0000-0000. Disponível em: <<https://sol.sbc.org.br/index.php/stil/article/view/17784>>.

HOREV, R. **BERT Explained: State of the art language model for NLP**. 2018. Published in Towards Data Science, Nov 10, 2018, 7 min read, 9.2K claps. Disponível em: <<https://towardsdatascience.com/bert-explained-state-of-the-art-language-model-for-nlp-f8b21a9b6270>>.

LIU, B. In: **Sentiment Analysis And Opinion Mining**. [S.l.]: Morgan Claypool, 2012.

LIU, Y. et al. **RoBERTa: A Robustly Optimized BERT Pretraining Approach**. 2019.

MALINGAN, N. **TensorFlow Word Embeddings**. 2024. Blog post. Disponível em: <<https://www.scaler.com/topics/tensorflow/tensorflow-word-embeddings/>>.

NARKHEDE, S. Understanding confusion matrix. **Towards Data Science**, May 2018. Disponível em: <<https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62>>.

PANG, B.; LEE, L. Opinion mining and sentiment analysis. **Foundations and Trends in Information Retrieval**, v. 2, n. 1-2, p. 1–135, 2008. Disponível em: <<http://www.cs.cornell.edu/home/llee/opinion-mining-sentiment-analysis-survey.html>>.

SHAKYA, M. M. S.; SHRESTHA, A. **Fine-grained Sentiment Classification using BERT**. 2019.

SOUZA, R. N. F.; LOTUFO, R. BERTimbau: pretrained BERT models for Brazilian Portuguese. In: **Proceedings of the Brazilian Conference on Intelligent Systems**. [S.l.]: Springer, 2020. p. 403–417.

TRIPATHI, R. **What are Vector Embeddings?** 2022. 27, 48, 49, 50, 52 p. Acesso: Setembro de 2022. Disponível em: <<https://www.pinecone.io/learn/vector-embeddings/>>.

VASWANI, A. et al. **Attention Is All You Need**. 2023.

WIEBE JANYCE M., B. R. F.; O'HARA, T. P. Development and use of a gold-standard data set for subjectivity classifications. In: **Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics**. College Park, Maryland, USA: Association for Computational Linguistics, 1999. p. 246–253. Disponível em: <<https://aclanthology.org/P99-1032>>.