

UNIVERSIDADE FEDERAL DE SÃO CARLOS
DEPARTAMENTO DE COMPUTAÇÃO
BACHARELADO EM ENGENHARIA DE COMPUTAÇÃO

Fernanda Malheiros Assi

**ANÁLISE DE EMOÇÕES EM TWEETS DE
RESPOSTA A POSTS DO EX-PRESIDENTE
DO BRASIL JAIR MESSIAS BOLSONARO**

São Carlos - SP

2023

Fernanda Malheiros Assi

**ANÁLISE DE EMOÇÕES EM TWEETS DE RESPOSTA
A POSTS DO EX-PRESIDENTE DO BRASIL JAIR
MESSIAS BOLSONARO**

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Engenharia de Computação da UNIVERSIDADE FEDERAL DE SÃO CARLOS, como requisito parcial para a obtenção do título de Bacharel em Engenharia de Computação.

Orientação Profa. Dra. Helena de Medeiros Caseli

São Carlos - SP

2023

Dedico este trabalho aos meus companheiros de viagem nesta jornada acadêmica, cuja amizade e apoio foram inestimáveis.

Agradecimentos

Gostaria de agradecer primeiramente aos meus pais, Sueli e Marcelo. Vocês foram o farol de luz constante em minha vida, guiando-me com seu amor incondicional e apoio. Aos meus avós, “Bazinha” e Caitano, minha gratidão pelo seu carinho e pelos exemplos de vida que sempre me inspiraram.

À Helena, Diego e Carol, professores e orientadores, sou imensamente grata por sua orientação, paciência e dedicação, por acreditarem em mim e por me ajudar com seus conhecimentos e mentoria. Sem vocês, nada disso seria possível.

Também gostaria de agradecer os meus amigos de faculdade, que estiveram ao meu lado durante todos os desafios e vitórias. Cada um de vocês tornou minha experiência universitária muito mais leve, com risadas, aprendizados e amizades inestimáveis.

À equipe do *Gate Team*, muito obrigada por me receber e acolher e pela valiosa base de dados que me foi fornecida por vocês, crucial para a realização deste trabalho.

Por fim, gostaria de agradecer à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pelo apoio financeiro com o projeto de Iniciação Científica nº 2021/07067-7 e com a Bolsa de Estágio em Pesquisa no Exterior (BEPE) nº 2022/02472-6.

*“O progresso é impossível sem mudança; e aqueles que não conseguem mudar suas mentes
não podem mudar nada.”
(George Bernard Shaw)*

Resumo

Nas últimas décadas, diversos avanços computacionais permitiram o acesso mais democrático à internet e, conseqüentemente, às redes sociais. O alcance da política digital é vasto e o papel do Twitter nesse cenário é inegável. Esta plataforma evoluiu para se tornar um grande palco para o debate político, onde cidadãos, líderes e instituições interagem diretamente. No Brasil, essa dinâmica é especialmente relevante dada a atividade intensa do ex-presidente Jair Messias Bolsonaro na plataforma. Entretanto, a análise manual de um volume tão grande de dados é impraticável e propensa a erros. Para que tais dados textuais possam ser transformados em informação, são necessárias tecnologias que consigam extrair e processar esses dados. O PLN oferece ferramentas valiosas para a análise automatizada de grandes volumes de texto. Dentre essas ferramentas, destaca-se o GoEmotions, um modelo de aprendizado de máquina que consegue identificar 27 categorias de emoções em textos. Desta forma, este trabalho utilizou o GoEmotions, adaptado para o português, para categorizar as emoções presentes em *tweets* de resposta aos *tweets* do ex-presidente Jair Bolsonaro, visando entender quais emoções predominam nesse contexto. A análise dos resultados mostrou um desequilíbrio substancial na presença das emoções no *cópus*, sendo a “raiva” a emoção mais predominante. Além disso, foi constatada uma performance variada do modelo na identificação das emoções, com maior precisão na identificação de “raiva”, mas apresentando desafios com emoções complexas como “admiração” e “curiosidade”, especialmente em contextos de ironia ou sarcasmo. Esses resultados apontam para o potencial dos modelos de PLN na análise de emoções em *tweets* políticos, ao mesmo tempo que destacam a necessidade de melhorias contínuas para lidar com nuances da linguagem.

Palavras-chave: Processamento de Linguagem Natural, Ciência Política Computacional, Detecção de Emoção, Política Brasileira, Twitter.

Abstract

In recent decades, several computational advances have allowed more democratic access to the internet and, consequently, to social networks. The reach of digital politics is vast, and Twitter's role in this scenario is undeniable. This platform became a significant political debate stage, where citizens, leaders, and institutions interact directly. In Brazil, this dynamic is especially relevant given the intense activity of former president Jair Messias Bolsonaro on the platform. However, manual analysis of such a large volume of data is impractical and error-prone. For such textual data to be transformed into information, it is necessary technologies to extract and process it. The NLP offers valuable tools for the automated analysis of large volumes of text. Among these tools, GoEmotions stands out, a machine learning model that can identify 27 categories of emotions in texts. Thus, this work used GoEmotions, adapted for Portuguese, to categorize the emotions in tweets in response to former president Jair Bolsonaro's tweets to understand which emotions predominate in this context. The analysis of the results showed a substantial imbalance in the presence of emotions in the corpus, with "anger" being the most predominant emotion. In addition, a varied performance of the model in the identification of emotions was verified, with greater precision in the identification of "anger", but presenting challenges with complex emotions such as "admiration" and "curiosity", especially in contexts of irony or sarcasm. These results point to the potential of NLP models in analyzing emotions in political *tweets* while also highlighting the need for continuous improvements to deal with language nuances.

Keywords: Natural Language Processing, Computational Political Science, Emotion Detection, Brazilian Politics, Twitter.

Lista de ilustrações

| | |
|--|----|
| Figura 1 – Rede Neural com camadas de entrada (<i>input layer</i>), oculta (<i>hidden layer</i>) e saída (<i>output layer</i>) | 29 |
| Figura 2 – Representação de um modelo <i>encoder-decoder</i> | 29 |
| Figura 3 – Representação da arquitetura de um <i>Transformer</i> | 33 |
| Figura 4 – Técnicas de classificação de sentimentos | 38 |
| Figura 5 – Espaço Semântico de Scherer | 40 |
| Figura 6 – Categorias de emoções ordenadas por frequência no dataset | 44 |
| Figura 7 – Distância Inter-tópicos | 45 |
| Figura 8 – Resultado dos modelos avaliados | 47 |
| Figura 9 – Resultados obtidos para as coletas referentes ao primeiro e segundo domingo eleitoral. | 48 |
| Figura 10 – Pipeline do projeto | 49 |
| Figura 11 – Painel do Kibana | 52 |
| Figura 12 – Resultados do GoEmotions | 56 |
| Figura 13 – Quantidade de instâncias por emoção | 57 |
| Figura 14 – <i>Boxplot</i> dos scores por emoção | 58 |
| Figura 15 – <i>Wordcloud</i> para Raiva | 59 |
| Figura 16 – <i>Wordcloud</i> para Curiosidade | 60 |
| Figura 17 – <i>Wordcloud</i> para Admiração | 60 |

Lista de tabelas

| | |
|--|----|
| Tabela 1 – Resultados das etapas de Pré-processamento | 54 |
| Tabela 2 – Palavras mais frequentes | 59 |
| Tabela 3 – Resultado da análise manual de 100 instâncias de cada classe (Raiva, Curiosidade, Admiração) | 61 |
| Tabela 4 – Exemplos de erros e acertos para Raiva | 62 |
| Tabela 5 – Exemplos de erros e acertos para Curiosidade | 64 |
| Tabela 6 – Exemplos de erros e acertos para Admiração | 64 |

Sumário

| | | |
|------------|--|-----------|
| 1 | INTRODUÇÃO | 19 |
| 1.1 | Motivação | 20 |
| 1.2 | Objetivos e Hipóteses | 21 |
| 1.3 | Organização do Trabalho | 21 |
| 2 | FUNDAMENTAÇÃO TEÓRICA | 23 |
| 2.1 | Inteligência Artificial | 23 |
| 2.2 | Aprendizado de Máquina | 24 |
| 2.2.1 | Aprendizado Supervisionado | 24 |
| 2.2.2 | Aprendizado Não-Supervisionado | 25 |
| 2.2.3 | Aprendizado Semi-Supervisionado | 25 |
| 2.3 | Processamento de Linguagem Natural | 25 |
| 2.3.1 | Pré-processamento de texto | 26 |
| 2.3.2 | Desenvolvimento de sistemas | 27 |
| 2.4 | Balanceamento de Classes | 27 |
| 2.5 | Redes Neurais | 28 |
| 2.5.1 | Modelos <i>Encoder-Decoder</i> | 28 |
| 2.6 | Modelos de Língua Neural | 30 |
| 2.6.1 | <i>Tokenização subword</i> | 30 |
| 2.6.2 | Mecanismos de Atenção | 31 |
| 2.6.3 | <i>Transformers</i> | 32 |
| 2.6.3.1 | BERT | 32 |
| 2.7 | Medidas de Avaliação | 34 |
| 3 | ANÁLISE DE SENTIMENTOS | 37 |
| 3.1 | Classificação de Emoções | 38 |
| 3.1.1 | Mapeamento de emoções em Espaços Afetivos | 39 |
| 3.1.2 | Taxonomia de emoções | 41 |
| 3.1.2.1 | Modelo de Emoções Básicas de Paul Ekman | 41 |
| 3.1.2.2 | GoEmotions | 42 |
| 3.2 | Detecção de emoções em <i>tweets</i> no contexto da política brasileira | 43 |
| 3.2.1 | Detecção de crenças de grupos | 44 |
| 3.2.2 | Detecção de polaridade no processo de <i>Impeachment</i> | 46 |
| 3.2.3 | Análise de Sentimentos nas Eleições de 2018 | 47 |
| 4 | MATERIAIS E MÉTODOS | 49 |

| | | |
|------------|--|-----------|
| 4.1 | Google Colab e Bibliotecas Python | 49 |
| 4.2 | corpus | 51 |
| 4.2.1 | Extração dos <i>tweets</i> | 51 |
| 4.2.1.1 | <i>Tweets</i> semente | 52 |
| 4.2.1.2 | Respostas aos <i>tweets</i> semente | 53 |
| 4.3 | Pré-processamento | 53 |
| 4.4 | GoEmotions para Português | 54 |
| 5 | RESULTADOS E DISCUSSÕES | 57 |
| 5.1 | Análise Exploratória | 57 |
| 5.2 | Análise de Erros | 60 |
| 5.2.1 | Raiva | 62 |
| 5.2.2 | Curiosidade | 63 |
| 5.2.3 | Admiração | 63 |
| 6 | CONCLUSÕES | 65 |
| 6.1 | Trabalhos Futuros | 65 |
| | REFERÊNCIAS | 67 |

1 Introdução

Nas últimas décadas, diversos avanços computacionais permitiram o acesso mais democrático à internet e, conseqüentemente, às redes sociais. Com isso, tem-se um aumento constante do número de dados acessíveis online e que estão disponíveis, em sua grande maioria, em formato de texto. Para que tais dados textuais possam ser transformados em informação, são necessárias tecnologias que consigam extrair e processar esses dados.

Desta forma, o Processamento de Linguagem Natural (PLN) é a subárea da Inteligência Artificial (IA) que combina conhecimentos da Ciência da Computação e Linguística com o objetivo de desenvolver sistemas inteligentes capazes de extrair, analisar e interpretar entradas em linguagem natural.

As aplicações de PLN estão presentes em diversas tecnologias de uso cotidiano, dentre elas, vale citar os tradutores automáticos de textos, como o Google Tradutor¹; programas que interagem com usuários, conhecidos como assistentes virtuais, como a Alexa da Amazon e o ChatGPT² (RADFORD et al., 2019); e os sistemas de recomendação, que tem por objetivo personalizar os conteúdos apresentados aos usuários e estão presentes em diversos sites, como os de *e-commerce* (SALUNKE; NICHITE, 2022) e em serviços de *streamming* (GOMEZ-URIBE; HUNT, 2016).

Além dos exemplos citados anteriormente, outra subárea do PLN de bastante interesse no momento é a Análise de Sentimento (AS), também conhecida como Mineração de Opinião (MO), que utiliza técnicas computacionais, de PLN e de Aprendizado de Máquina (AM) para extrair sentimentos em textos de linguagem natural. Ela é mais comumente usada para determinar a valência (ou polaridade) de um texto, que pode ser positiva, negativa ou neutra. No entanto, também existem classificadores focados na detecção de emoções humanas em textos, entre eles, pode-se citar o GoEmotions (DEMSZKY et al., 2020).

Diversas empresas aplicam as técnicas de AS em *reviews* de seus produtos e usufruem da saída dos classificadores para ajudar em tomadas de decisões. A AS também pode ser utilizada em diversos outros contextos, como em redes sociais, sites de notícias e no mercado de ações. No mercado de ações, a AS é aplicada em sites de notícias para, em conjunto com técnicas de previsão de séries temporais, tentar prever valores de ações futuros (BHARATHI; GEETHA, 2017).

Uma outra aplicação importante das técnicas de AS é no cenário político. Com a ascensão das mídias sociais, os políticos têm a oportunidade de se engajar diretamente

¹ <<https://translate.google.com.br/>>

² <<https://chat.openai.com/>>

com o público. Através da AS, é possível medir a resposta a esses engajamentos, identificar tendências de sentimentos e reações a eventos ou políticas específicas (TUMASJAN et al., 2010). As informações obtidas podem ser usadas para moldar estratégias de comunicação, políticas e tomada de decisões. É importante notar que, embora a AS e a MO ofereçam muitas oportunidades no campo da política, também há desafios, como a detecção de ironia e sarcasmo (REYES; ROSSO; VEALE, 2013) e a distinção entre opiniões verdadeiras e desinformação (SHU et al., 2017).

As redes sociais, particularmente o Twitter³, tornaram-se um palco global para o debate político, onde cidadãos, líderes e instituições interagem diretamente. No Brasil, essa dinâmica é especialmente relevante dada a atividade intensa do ex-presidente Jair Messias Bolsonaro na plataforma. Assim, este trabalho tem por objetivo analisar das reações dos *tweets* do Bolsonaro no período da pandemia da COVID-19, o que pode oferecer um panorama singular das emoções da população frente ao seu mandato.

Entretanto, a análise manual de um volume tão grande de dados é impraticável e propensa a erros. O PLN, oferece ferramentas valiosas para a análise automatizada de grandes volumes de texto. Dentre essas ferramentas, destaca-se o GoEmotions, um modelo de aprendizado de máquina que consegue identificar 27 categorias de emoções em textos. Desta forma, este trabalho visa utilizar o GoEmotions, adaptado para o português (HAMMES; FREITAS, 2021), para categorizar as emoções presentes em *tweets* de resposta aos *tweets* do ex-presidente Jair Bolsonaro, visando entender quais emoções predominam nesse contexto.

1.1 Motivação

Uma consequência da popularização da internet e de seu fácil acesso, as redes sociais vêm se tornando, cada vez mais, um espaço onde as pessoas compartilham suas opiniões, sentimentos e percepções sobre diversos assuntos, tanto pessoais quanto culturais e políticos. Neste contexto, diversos trabalhos de processamento de língua natural (PLN) têm focado no processamento de textos em mídias sociais para, entre outros, prever ideologias políticas (IYYER et al., 2014), transtornos de personalidade (PREOTIUC-PIETRO et al., 2016) e saúde mental (COPPERSMITH; DREDZE; HARMAN, 2014).

Este trabalho foi desenvolvido no contexto da pandemia do COVID-19, que impactou significativamente as áreas de saúde, econômica, social e ambiental em todo o mundo (MOFIJUR et al., 2021). Em um contexto de tantas incertezas, discursos de líderes mundiais, mesmo em contextos informais, podem impactar consideravelmente o comportamento da população e o cenário político, econômico e social (AJZENMAN; CAVALCANTI;

³ <<https://twitter.com/>>

MATA, 2020). No entanto, os discursos do ex-presidente do Brasil, Jair Messias Bolsonaro, durante a pandemia foram repletos de polêmica e desinformação (KIBUUKA, 2020).

O alcance da política digital é vasto e o papel do Twitter nesse cenário é inegável. O registro digital das reações à presidência de Bolsonaro, em forma de *tweets*, é uma valiosa fonte de dados para entender o clima político e emocional da época. Ao explorar esses dados, espera-se contribuir para uma maior compreensão do impacto da presidência de Bolsonaro na esfera emocional da política brasileira, além de destacar o valor das ferramentas de PLN na análise do discurso político.

1.2 Objetivos e Hipóteses

Este trabalho tem por objetivo analisar as emoções expressas em *tweets* relacionados ao contexto da política brasileira, mais especificamente aos *tweets* de resposta aos *tweets* do ex-presidente do Brasil, Jair Messias Bolsonaro. Para isso, o GoEmotions para Português será utilizado como ferramenta de rotulação de emoções.

Além disso, a avaliação da qualidade da saída do GoEmotions neste contexto específico também é um objetivo secundário importante. O desenvolvimento de uma análise exploratória dos dados e das emoções mais frequentes permite gerar uma imagem detalhada das reações emocionais ao ex-presidente durante o período considerado. Ao mesmo tempo, a avaliação da performance do GoEmotions nesse contexto fornece um importante ponto de partida para futuras pesquisas que buscam aplicar ferramentas similares na análise do discurso político nas redes sociais.

As hipóteses baseiam-se na suposição de que o GoEmotions para Português será uma ferramenta adequada para identificar diferenças significativas nas emoções expressas no contexto político brasileiro em *tweets*. Com isso, espera-se que seja possível fornecer uma melhor compreensão sobre a atual atmosfera política brasileira.

1.3 Organização do Trabalho

Este trabalho está dividido em seis capítulos:

- No Capítulo 1 são apresentados a proposta do trabalho, a motivação, os objetivos e hipóteses da pesquisa;
- O Capítulo 2 compreende a fundamentação teórica necessária para a realização da pesquisa, passando por diversos conceitos de Aprendizado de Máquina e Processamento de Linguagem Natural e formas de análise de resultados.

- O Capítulo 3 apresenta os principais trabalhos relacionados a Análise de Sentimento e Classificação de Emoções que ajudaram a embasar e fundamentar a pesquisa;
- O Capítulo 4 descreve os materiais e os métodos utilizados para o desenvolvimento deste trabalho, além de mostrar o fluxo de desenvolvimento da pesquisa;
- O Capítulo 5 apresenta uma análise dos resultados encontrados;
- Finalmente, no Capítulo 6, são apresentadas as conclusões obtidas no trabalho e as sugestões para pesquisas futuras.

2 Fundamentação Teórica

Este Capítulo apresenta os principais conceitos que embasam o presente trabalho. São tratados dos conceitos de Inteligência Artificial (Seção 2.1), Aprendizado de Máquina (Seção 2.2), Processamento de Linguagem Natural (Seção 2.3), Balanceamento de Classes (Seção 2.4), Modelos de Redes Neurais (Seção 2.5) e Modelos de Língua Neurais (Seção 2.6). Por fim são discutidas as principais medidas de avaliação utilizadas para modelos de Aprendizado de Máquina (Seção 2.7).

2.1 Inteligência Artificial

A Inteligência Artificial (IA), termo cunhado por John McCarthy em 1956, é um dos campos de estudo da Ciência da Computação, e pode ser definida como “*A simulação da inteligência humana por sistemas computacionais.*” (RUSSELL; NORVIG, 2009). Essa definição destaca que a IA trata da criação de sistemas computacionais capazes de imitar ou emular habilidades cognitivas humanas, como aprendizado, resolução de problemas e tomada de decisões.

Além disso, espera-se que a IA seja capaz de aprender e se transformar ao longo do tempo, fazendo melhorias em seu desempenho à medida que processa mais dados. Em outras palavras, uma IA precisa possuir capacidade de adaptação. Desta forma, os algoritmos de IA almejam encontrar, de forma autônoma, a solução ótima de um determinado problema. Vale ressaltar que, portanto, as soluções baseadas em IA não são universais, e apresentam particularidades em cada problema.

A área de IA é a maior subárea da computação que este trabalho abrange e é a base para inúmeros campos de estudos e aplicações diversas. Dentro do contexto da IA, alguns exemplos de aplicações incluem:

- **Aprendizado de Máquina:** desenvolvimento de algoritmos e modelos estatísticos que permitem que as máquinas aprendam com os dados e façam previsões ou tomem decisões sem serem explicitamente programadas para isso;
- **Processamento de Linguagem Natural (PLN):** uso de IA para entender, interpretar e gerar linguagem humana;
- **Visão Computacional:** uso de IA para processar e analisar imagens e vídeos, permitindo que as máquinas entendam e interpretem informações visuais;
- **Robótica:** utilização da IA para criação de robôs mais autônomos.

2.2 Aprendizado de Máquina

O Aprendizado de Máquina (AM) é um campo de estudo da IA que se concentra no desenvolvimento de algoritmos e modelos estatísticos que permitem que os computadores aprendam e façam previsões ou tomem decisões com base em dados, sem que sejam explicitamente programados para a tarefa (ALPAYDIN, 2010). A base do AM está na ideia de que um sistema de computador pode aprender com dados, identificar padrões e relacionamentos e fazer previsões com base nesses padrões. Existem vários tipos de algoritmos de AM, cada um projetado para lidar com diferentes tipos de problemas e dados.

As técnicas de AM têm sua base no aprendizado indutivo. Nele, a aprendizagem de um conceito ocorre através da inferência indutiva feita sobre os exemplos apresentados. Isso significa que as hipóteses geradas por meio desta inferência podem representar, ou não, a verdade (REZENDE, 2003). Assim, o algoritmo é treinado com conjunto de dados de treinamento e, em seguida, usa esse conhecimento para fazer previsões ou generalizações sobre dados novos e não vistos.

O aprendizado indutivo e, portanto, o AM, são subdivididos em três principais categorias, conforme o grau de supervisão: Aprendizado Supervisionado, Não Supervisionado e Semi-Supervisionado. As subseções seguintes introduzem cada uma dessas subdivisões.

2.2.1 Aprendizado Supervisionado

Os algoritmos de aprendizado supervisionado aprendem com dados de treinamento rotulados por humanos para fazer previsões sobre dados novos e não rotulados. Para isso, o algoritmo recebe um conjunto de pares de entrada e saída e o objetivo é aprender a relação entre as entradas e as saídas. O modelo computacional resultante desse treinamento, então, usa essa relação aprendida para fazer previsões sobre a saída para novas entradas.

O Aprendizado Supervisionado possui diversos tipos de algoritmos, como os de classificação e regressão. Os problemas de regressão envolvem a previsão de um valor de saída contínuo, como o preço de um estoque ou a probabilidade de um cliente comprar determinado produto. Os problemas de classificação, por outro lado, envolvem a atribuição de uma classe ou rótulo específico, como classificar um e-mail como spam ou não spam. Assim, a escolha do algoritmo depende do tipo de problema e da natureza dos dados. No entanto, algoritmos como árvores de decisão (WU et al., 2008) e florestas aleatórias (HO, 1995) podem ser utilizados tanto para problemas de regressão quanto para problemas de classificação.

2.2.2 Aprendizado Não-Supervisionado

Muitas vezes, não há disponível dados rotulados, desta forma, uma nova abordagem precisa ser considerada. Com isso, os algoritmos de aprendizado não-supervisionado aprendem com dados não rotulados para descobrir padrões e relacionamentos ocultos nos dados. Ao contrário do aprendizado supervisionado, o aprendizado não-supervisionado não tem nenhuma saída específica para prever. Em vez disso, o objetivo é identificar padrões e estruturas nos dados, como agrupar pontos de dados semelhantes ou identificar subgrupos subjacentes nos dados.

Os algoritmos de aprendizado não-supervisionado incluem agrupamento k-means (LLOYD, 1982), e técnicas de redução de dimensionalidade, como análise de componentes principais (PCA) (F.R.S., 1901). Esses algoritmos ajudam a encontrar estrutura nos dados e fornecem uma compreensão mais intuitiva das relações presente nos dados. Uma das principais vantagens do aprendizado não-supervisionado é que ele pode lidar com grandes quantidades de dados não rotulados, e identificar padrões e relações que não seriam facilmente perceptíveis de outra forma. No entanto, muitas vezes os resultados podem ser difíceis de serem interpretados e existe uma maior dificuldade em medir a eficácia do algoritmo, visto que não existem dados rotulados (DOMINGOS, 2012).

2.2.3 Aprendizado Semi-Supervisionado

O aprendizado semi-supervisionado combina o aprendizado supervisionado e o não-supervisionado. Nele, os algoritmos realizam o treinamento em uma combinação de dados rotulados e não rotulados. Os dados rotulados fornecem ao algoritmo as saídas para entradas específicas, enquanto os dados não rotulados fornecem informações adicionais para ajudar o algoritmo a aprender mais sobre a estrutura e os relacionamentos nos dados. Essa abordagem pode ser particularmente útil quando há uma quantidade limitada de dados rotulados disponíveis, pois permite que o algoritmo faça uso de informações adicionais dos dados não rotulados para melhorar suas previsões (ZHU; GOLDBERG, 2009).

2.3 Processamento de Linguagem Natural

O Processamento de Linguagem Natural (PLN) é uma subárea da IA que tem seu principal foco na interação entre os computadores e a linguagem humana. O principal objetivo do PLN é desenvolver sistemas inteligentes capazes de extrair, analisar e interpretar entradas em linguagem natural (NADKARNI; OHNO-MACHADO; CHAPMAN, 2011). Para isso, são utilizados conhecimentos de computação e linguística para a criação de novas técnicas de PLN.

O fluxo de desenvolvimento de algoritmos de PLN ocorre em duas etapas principais:

o pré-processamento de dados e o desenvolvimento de algoritmos. O processamento de dados tem por objetivos preparar os dados de texto para serem analisados pelas máquinas. Uma vez que os dados foram limpos, começa a etapa de desenvolvimento de algoritmos, para que esses dados possam ser processados.

As aplicações de PLN são inúmeras, algumas das quais são: Análise de Sentimentos, Tradução Automática, Reconhecimento de Entidade Nomeada, Reconhecimento de Voz, Classificação e Categorização de Texto, Sistemas de Recomendação e Programas de Pergunta-Resposta. A Seção ?? abordará a aplicação de Análise de Sentimento.

2.3.1 Pré-processamento de texto

Muitas vezes, dados em formato de texto contém ruídos, como erros de grafia, pontuação e símbolos, que podem atrapalhar os algoritmos de AM. Assim, na primeira etapa no pipeline desenvolvimento de modelos de PLN é realizado o pré-processamento de texto. O objetivo é transformar o texto para um formato que pode ser facilmente analisado e processado por algoritmos de AM. O pré-processamento de texto geralmente inclui várias etapas, como:

- **Letras minúsculas:** converter todo o texto em letras minúsculas para garantir a consistência do texto e remover a distinção entre letras maiúsculas e minúsculas.
- **Remoção de pontuação:** Remoção de todos os sinais de pontuação do texto para simplificá-lo e otimizar seu processamento.
- **Tokenização:** quebrar o texto em unidades menores, como palavras ou frases individuais, chamados de *tokens*. Esta é uma etapa importante, visto que a maneira mais comum de processar o texto bruto ocorre no nível do *token*.
- **Remoção *stopwords*:** remoção de palavras muito frequentes ou que não agregam significativamente ao texto, e podem afetar os resultados da análise.
- **Stematização e Lematização:** Reduzir as palavras à sua forma raiz, diminuindo a dimensionalidade dos dados.
- **Remoção de caracteres especiais:** Remoção de quaisquer caracteres especiais, como números e símbolos, que podem ser considerados como ruído para os modelos de AM.

É válido ressaltar que não há um pré-processamento universal para todos os problemas, e a escolha de quais etapas usar está diretamente ligada aos objetivos da tarefa em questão e também depende de outros fatores como o tipo de dados e a capacidade de processamento computacional.

2.3.2 Desenvolvimento de sistemas

A etapa de desenvolvimento de algoritmos ocorre após o pré-processamento do texto, e é responsável por produzir programas capazes de gerar, analisar e interpretar a linguagem humana. Existem diversos tipos diferentes de sistemas e algoritmos de PLN, e o uso de cada um depende do tipo de problema e dos objetivos esperados. Os dois sistemas mais comumente usados são:

- **Sistema baseado em regras:** utiliza um conjunto de regras linguísticas e um mecanismo de inferência para tomar decisões. As regras são especificadas como declarações. Os sistemas baseados em regras eram usados no começo no desenvolvimento do PLN. No entanto, eles ainda são comumente usados em domínios em que a experiência humana é necessária, como diagnóstico médico (GAYATHRI; GOPICHAND; AMALI, 2018), análise financeira (MAYER et al., 2020) e atendimento ao cliente (MNASRI, 2019).
- **Sistema baseado em AM:** utilizam conhecimentos de estatística para o desenvolvimento de modelos de AM. Os sistemas baseados em AM são utilizados em diversas situações, como quando se tem um grande volume de dados, quando a tarefa possui alta complexidade, ou em casos onde há muito dados ambíguos.

2.4 Balanceamento de Classes

Em conjuntos de dados do mundo real, é comum que a distribuição de classes esteja altamente desequilibrada, com uma ou mais classes possuindo um número muito menor de instâncias do que as outras classes. O desbalanceamento dos dados pode fazer com que os modelos de AM tenham mais dificuldade em aprender a tarefas a partir dos dados, levando a previsões tendenciosas e menor desempenho nas avaliações (HE; GARCIA, 2009). Abordar o desbalanceamento é, portanto, uma consideração importante no projetos de AM, e várias técnicas foram desenvolvidas para ajudar a reduzir o impacto do desbalanceamento dos dados, dentre elas:

- **Oversampling:** geração de dados sintéticos para a classe minoritária para aumentar sua representação no conjunto de dados. Uma técnica de *oversampling* bastante utilizadas é o SMOTE (*Synthetic Minority Oversampling Technique*) (CHAWLA et al., 2002).
- **Undersampling:** reduzir o número de instâncias da classe majoritária para equilibrar a distribuição das classes. Usualmente as instâncias removidas são escolhidas de maneira aleatória.

- **Peso para as classes:** atribuição de pesos diferentes para as classes durante o treinamento para equilibrar sua contribuição para a perda geral. Um exemplo é o método *Class Balance Loss* (CUI et al., 2019), utilizado no GoEmotions para Português (HAMMES; FREITAS, 2021).

2.5 Redes Neurais

Redes Neurais são um tipo de modelo de AM inspirado na estrutura e funcionamento do cérebro humano (LUGER; STUBBLEFIELD, 1997). As redes neurais consistem em unidades de processamento interconectadas chamadas neurônios artificiais que são organizadas em camadas. Cada neurônio artificial recebe e processa as entradas e emite um sinal que é transmitido para a próxima camada de neurônios. As entradas são tipicamente representadas como valores numéricos, e o processamento realizado por cada neurônio é determinado por um conjunto de parâmetros ajustáveis, conhecidos como pesos, que são aprendidos pelo modelo durante o treinamento (HAYKIN, 2009).

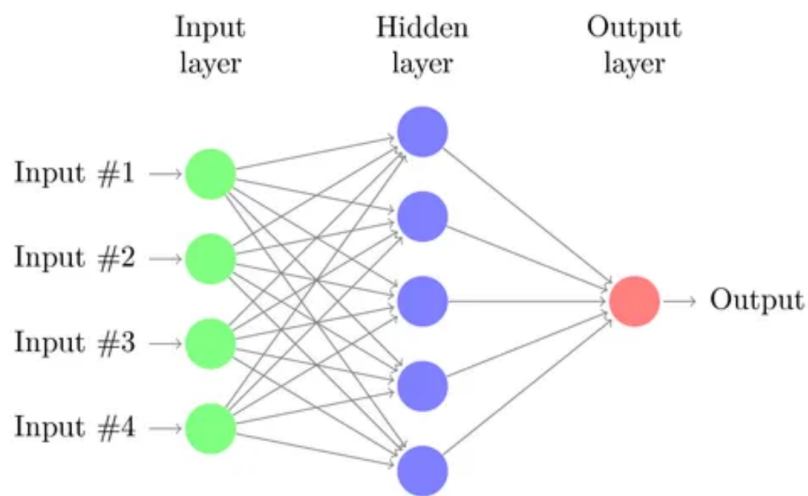
As redes neurais são tipicamente formadas por múltiplas camadas de neurônios artificiais. A Figura 1 mostra um exemplo esquemático de uma rede neural com três camadas: de entrada (*input layer*), oculta (*hidden layer*) e de saída (*output labels*). Cada camada possui um propósito específico:

- **Camada de entrada:** é a primeira camada da rede neural e seus neurônios recebem os dados de entrada. A camada de entrada normalmente não realiza nenhum cálculo, e apenas passa os dados de entrada para a próxima camada.
- **Camada oculta:** são as camadas intermediárias na rede neural e realizam a computação que transforma os dados de entrada nos dados de saída. Além disso, as camadas ocultas podem consistir em diferentes tipos de neurônios, dependendo do tipo de rede que está sendo usada.
- **Camada de saída:** é a última camada da rede neural e seus neurônios produzem a saída final do modelo. A camada de saída consiste em um único neurônio para problemas de regressão e em múltiplos neurônios para problemas de classificação, onde cada neurônio representa uma classe diferente.

2.5.1 Modelos *Encoder-Decoder*

Modelos *Encoder-Decoder* é um tipo de arquitetura de redes neurais amplamente utilizado, em especial nas tarefas de PLN e Visão Computacional. Além da camada oculta, os modelos possuem mais duas camadas: o *encoder*, ou codificador e o *decoder*, ou decodificador. O *encoder* é responsável por codificar a sequência de entrada em um vetor

Figura 1 – Rede Neural com camadas de entrada (*input layer*), oculta (*hidden layer*) e saída (*output layer*)

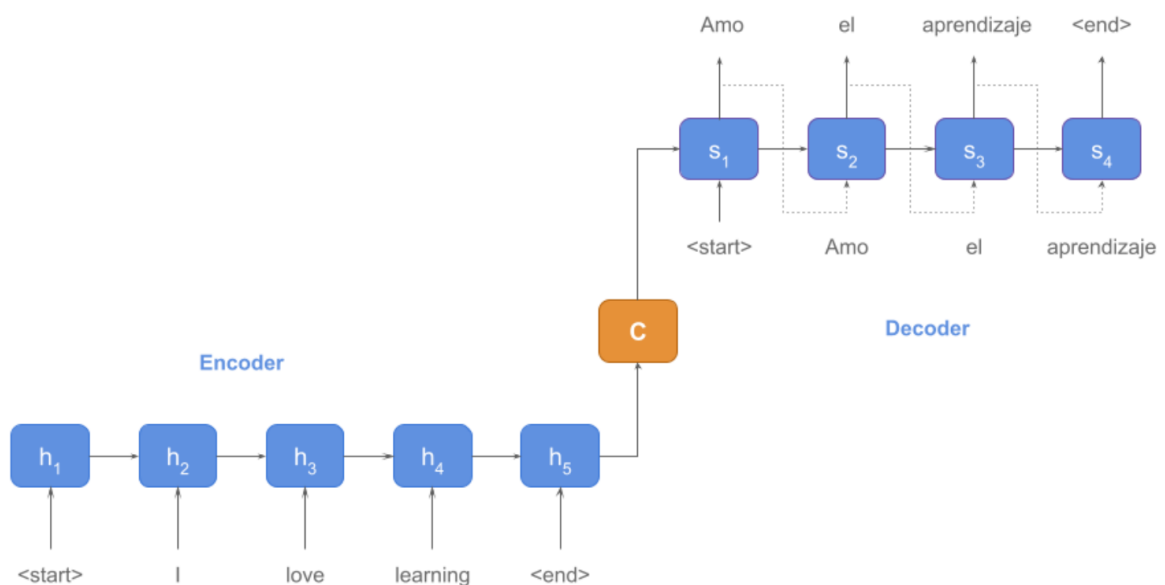


Fonte: (GUDIKANDULA, 2019)

de tamanho fixo. Já o *decoder*, decodifica o vetor na saída do modelo, que pode estar em diferentes formatos, como textos, rótulos e *tags* (SUTSKEVER; VINYALS; LE, 2014).

A Figura 2 mostra uma representação de um modelo *encoder-decoder* para o caso da tradução automática, onde o bloco laranja *C* representa as camadas ocultas.

Figura 2 – Representação de um modelo *encoder-decoder*



Fonte: (IBÁNEZ, 2022)

2.6 Modelos de Língua Neural

Os modelos de linguagem neural são um tipo de modelo estatístico e de linguagem baseados em aprendizado profundo que usam redes neurais artificiais para prever a probabilidade de uma sequência de palavras em um determinado idioma (BENGIO et al., 2003). O objetivo desses modelos é aprender a estrutura e os padrões da linguagem, que podem ser usados para diversas tarefas de PLN, dentre elas:

- **Análise de sentimento:** envolve prever o sentimento de um determinado trecho de texto, sendo mais comumente utilizado para previsão de polaridade (sentimento positivo, negativo ou neutro) (ZHANG; LIU, 2012), embora também existam trabalhos que usufruem de um conjunto maior de emoções (como raiva, alegria, entre outras) para realizar previsões (LIU et al., 2023). Esses modelos são treinados em dados de texto anotados e podem ser usados para analisar o sentimento de avaliações, postagens de mídia social ou outros tipos de dados de texto.
- **Geração de texto:** podem ser usados para gerar um novo texto com base nos padrões que aprenderam com os dados de treinamento. Os modelos podem ser usados em diversos contextos, como para gerar respostas realistas em sistemas de diálogo, escrever poesia ou ficção, ou gerar artigos de notícias (RADFORD et al., 2019).
- **Tradução automática:** traduzir texto de um idioma para outro. Os modelos são treinados em dados de texto paralelo, que consistem em sentenças em um idioma e suas traduções correspondentes em outro idioma (BAHDANAU; CHO; BENGIO, 2016).

Existem vários componentes dos modelos de linguagem neural que contribuíram para seu avanço, incluindo a *tokenização subword*, os mecanismos de atenção e os *transformers*. As subseções seguintes explicam brevemente esses componentes.

2.6.1 *Tokenização subword*

A *Tokenização subword* é uma técnica amplamente utilizada em modelos de linguagem natural e em outras tarefas de PLN. Nessa técnica de tokenização, ao invés de processar as palavras como unidades inteiras, elas são quebradas e processadas em unidades menores, conhecidas como *subwords* (ou subpalavras, em português) (BOJANOWSKI et al., 2017).

Isso ocorre visto que muitos modelos de PLN são treinados com conjunto de dados de dimensões muito grandes, o que faz com que o número de parâmetros do modelo também seja grande. A *Tokenização subword* é, então, usada para reduzir o vocabulário. Assim, palavras mais frequentes são mantidas inteiras e as palavras que ocorrem menos são

divididas em subpalavras, que muitas vezes já estão presentes no vocabulário. Por exemplo, a palavra “*refatorando*” pode ser dividida em três subpalavras: o prefixo “*re*”, “*fator*”, e o sufixo “*ando*”. É provável que cada subpalavra ocorra com mais frequência no texto do que a palavra original, visto que os prefixo “*re*” e o sufixo “*ando*” também possuem ocorrência em diversas outras palavras. Assim, o significado do texto é mantido, e o vocabulário é reduzido.

Existem diversos métodos de *Tokenização subword*, dentre eles, os principais são:

- **Byte-Pair Encoding (BPE)**: combina iterativamente o par de caracteres mais frequente no texto em uma única unidade até que o tamanho de vocabulário desejado seja alcançado. Isso resulta em um conjunto de *subwords* que capturam as semelhanças morfológicas e fonéticas das palavras, em vez de serem limitadas por limites lexicais estritos (SENNRICH; HADDOW; BIRCH, 2015).
- **WordPiece**: é semelhante ao BPE, no entanto, ao invés de usar os caracteres mais frequentes, o par de caracteres que maximiza a probabilidade dos dados de treinamento são combinados iterativamente no texto em uma única unidade, até que o tamanho do vocabulário desejado seja alcançado (WU et al., 2016). Esse é o tipo de tokenização utilizada nos modelos BERT (Seção 2.6.3.1).

2.6.2 Mecanismos de Atenção

O mecanismo de atenção (BAHDANAU; CHO; BENGIO, 2014) é um componente fundamental de vários modelos de língua. Ele surgiu com o objetivo de melhorar a performance do modelo *encoder-decoder* (Seção 2.5.1) para a tarefa de Tradução Automática. As redes neurais geralmente processam todos os dados de entrada de maneira sequencial ou paralela, sem a existência de nenhum mecanismo que determine a importância de cada elemento. Isso é especialmente problemático para sequências longas, assim os mecanismos de atenção surgiram para resolver este problema.

O mecanismo de atenção permite que o modelo atribua um peso ou pontuação a cada elemento de entrada, indicando sua importância para fazer uma previsão. Os elementos de entrada ponderados são, então, combinados para gerar um vetor de contexto, que é usado para fazer a previsão final. Dessa forma, o mecanismo de atenção permite que o modelo se concentre dinamicamente em diferentes partes dos dados de entrada conforme necessário, em vez de tratar todos os elementos igualmente.

Existem diferentes tipos de mecanismos de atenção, cada um com suas vantagens e desvantagens. Dentre eles, é válido citar:

- **Atenção Generalizada**: compara cada elemento da sequência de entrada com a sequência de saída. Isto quer dizer que, em cada iteração o *encoder* captura a

sequência de entrada e a compara com cada elemento do *decoder*. Com base nas pontuações de comparação, o mecanismo consegue selecionar as palavras às quais precisa prestar mais atenção (PANDYA, 2022).

- **Auto-Atenção:** pega partes específicas em diferentes posições na sequência de entrada e, ao longo do tempo, calcula uma composição inicial da sequência de saída. Ele não leva em consideração a sequência de saída, pois não há procedimento de entrada manual de dados em que a previsão da sequência de saída seja auxiliada de alguma forma. Os *transformers* (VASWANI et al., 2017) tem como base esse tipo de mecanismo de atenção.

2.6.3 Transformers

Um *Transformer* (VASWANI et al., 2017) é um tipo de modelo de língua neural baseado em mecanismos de auto-atenção e em camada de *encoders-decoders*. A Figura 3 mostra a representação da arquitetura de um *transformer*, com seis *encoders* e seis *decoders*, conforme inicialmente proposto. Cada *encoder* é formado por uma camada de auto-atenção e por uma rede *Feed Forward*. Já o *decoder*, além dessas duas camadas, também possui a camada *decoder attention*, que é a camada de atenção do decodificador. O primeiro *encoder* recebe o vetor de entrada e os processa antes de passar para o próximo *encoder*. Quando o último *encoder* terminar de processar a entrada, ele passa a informação para o primeiro *decoder* e o processo se repete.

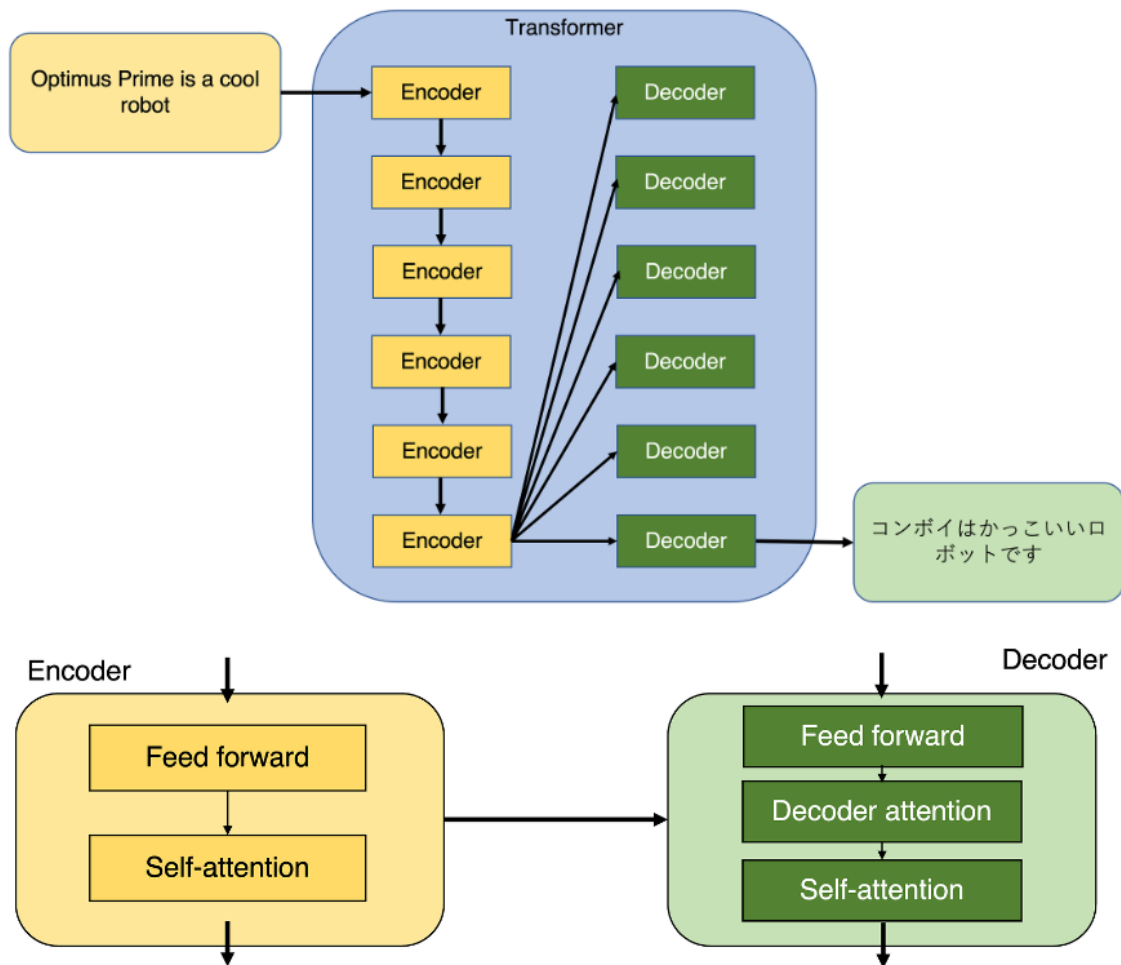
Uma vantagem importante da arquitetura do *transformer* em relação a outras redes neurais, é que a presença de contexto em torno de uma palavra em uma distância grande é feita de maneira mais eficiente. Além disso, ao contrário de alguns modelos de redes neurais, os dados de entrada não precisam ser processados sequencialmente, o que permite a paralelização dos dados. Os *Transformers* têm se tornado a ferramenta base para diversos modelos de PLN, como o BERT (DEVLIN et al., 2019) (Seção 2.6.3.1) e o GPT-3 (BROWN et al., 2020).

2.6.3.1 BERT

O BERT (*Bidirectional Encoder Representations from Transformers*) (DEVLIN et al., 2019) é um modelo de aprendizado profundo com a arquitetura baseada em *Transformers*, desenvolvido pela Google em 2018. O BERT é utilizado em inúmeras tarefas de PLN, como classificação de textos, análise de sentimentos (como o GoEmotions, discutido na Seção 3.1.2.2) e tradução automática, tendo obtido estado-da-arte em grande parte delas.

Ao contrário dos modelos direcionais, o BERT pode ser considerado como um modelo bidirecional. Isso significa que, ao invés de processar a entrada de texto sequencialmente

Figura 3 – Representação da arquitetura de um *Transformer*



Fonte: (IBÁNEZ, 2022)

(da esquerda para direita ou da direita para a esquerda), ele lê a sequência inteira simultaneamente. Essa característica faz com que o BERT seja capaz de capturar o contexto de uma palavra, baseada em todas as palavras ao seu entorno, o que não é possível em modelos unidirecionais.

Além disso, é possível realizar ajuste fino dos modelos para tarefas específicas de PLN. Isso é feito adicionando algumas camadas sobre o modelo BERT pré-treinado e treinando em um conjunto de dados rotulados. Existem diversas variantes do modelo BERT, dentre elas, destacam-se o *BERT-base*, o modelo base do BERT, o *BERT-Large*, uma versão mais complexa do BERT, com mais blocos de *transformers* e parâmetros e o *BERTimbau* (SOUZA; NOGUEIRA; LOTUFO, 2020), um modelo BERT pré-treinado para o Português Brasileiro.

2.7 Medidas de Avaliação

As medidas de avaliação são métricas numéricas usadas para avaliar o desempenho de um determinado modelo. Elas fornecem uma maneira de quantificar o desempenho do modelo na solução de um problema específico. Muitas vezes os resultados das avaliações são comparados com os de outros modelos, assim é possível obter seu desempenho relativo e determinar qual modelo é mais adequado para o problema (JAPKOWICZ; SHAH, 2011).

A escolha da métrica de avaliação depende do problema e do tipo de modelo que está sendo usado. É importante escolher a métrica de avaliação apropriada que se alinhe aos objetivos do modelo e ao problema que está sendo resolvido. Geralmente, nos modelos de AM supervisionados, uma parte do conjunto de dados, chamada de conjunto de teste, é separada para avaliar o modelo, não sendo usada em seu treinamento. Existem inúmeras métricas que podem ser utilizadas para avaliar modelos de AM supervisionados, a seguir, serão citados alguns exemplos das técnicas mais comumente utilizadas.

Nas Equações 2.1, 2.2, 2.3 e 2.4, VP (Verdadeiro Positivo) indica a quantidade de instâncias para as quais o modelo predisse de forma correta a classe, analogamente, VN (Verdadeiro Negativo), é a quantidade de instâncias para as quais o modelo predisse de forma correta uma classe negativa. O FP (Falso Positivo) acontece quando o modelo prediz de forma errada uma classe positiva. Por fim, o FN (Falso Negativo) ocorre quando o modelo prediz incorretamente uma classe negativa.

- **Acurácia:** mede a proporção dos resultados com predição correta em relação a todas as predições. A Equação 2.1 mostra a fórmula da acurácia. Ela fornece uma compreensão geral do desempenho de um modelo, no entanto, pode ser ineficiente quando o conjunto de dados está desbalanceado.

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN} \quad (2.1)$$

- **Precisão:** mede a proporção dos resultados corretamente preditos como positivos sobre todas as predições rotuladas como positivas. Em outras palavras, a precisão visa responder a pergunta: “De todos os dados classificados como positivos, quantos são realmente positivos?”. A Equação 2.2 mostra a fórmula da precisão

$$Precisão = \frac{VP}{VP + FP} \quad (2.2)$$

- **Recall:** também é conhecido como cobertura, e mede a proporção dos resultados corretamente preditos como positivos sobre todas as instâncias reais positivas. Assim, dá uma ideia da completude das previsões positivas feitas pelo modelo. A Equação 2.3 mostra a fórmula do *recall*.

$$Recall = \frac{VP}{VP + FN} \quad (2.3)$$

- **F1-score** combina a precisão e o *recall* em uma única métrica, com o objetivo de trazer ser uma métrica única que determina a qualidade geral do modelo. Ela definida como a média harmônica de precisão e recall, e sua fórmula está presente na Equação 2.4.

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

3 Análise de Sentimentos

Neste capítulo, serão abordados os principais conceitos relacionados à Análise de Sentimento (AS) e à Classificação de Emoções em textos (Seção 3.1). A discussão teórica se concentrará em fornecer um panorama abrangente dessas áreas, elucidando os métodos e técnicas comumente empregados para a extração e categorização de emoções e sentimentos expressos em linguagem natural. Além disso, serão apresentados trabalhos relacionados que focam especificamente na detecção e análise de emoções em *tweets* no contexto na política brasileira (Seção 3.2).

A Análise de Sentimento (AS), ou Mineração de Opinião (MO) é uma subárea do PLN que utiliza técnicas computacionais, de PLN e de Aprendizado de Máquina (AM), para extrair sentimentos em textos de linguagem natural.

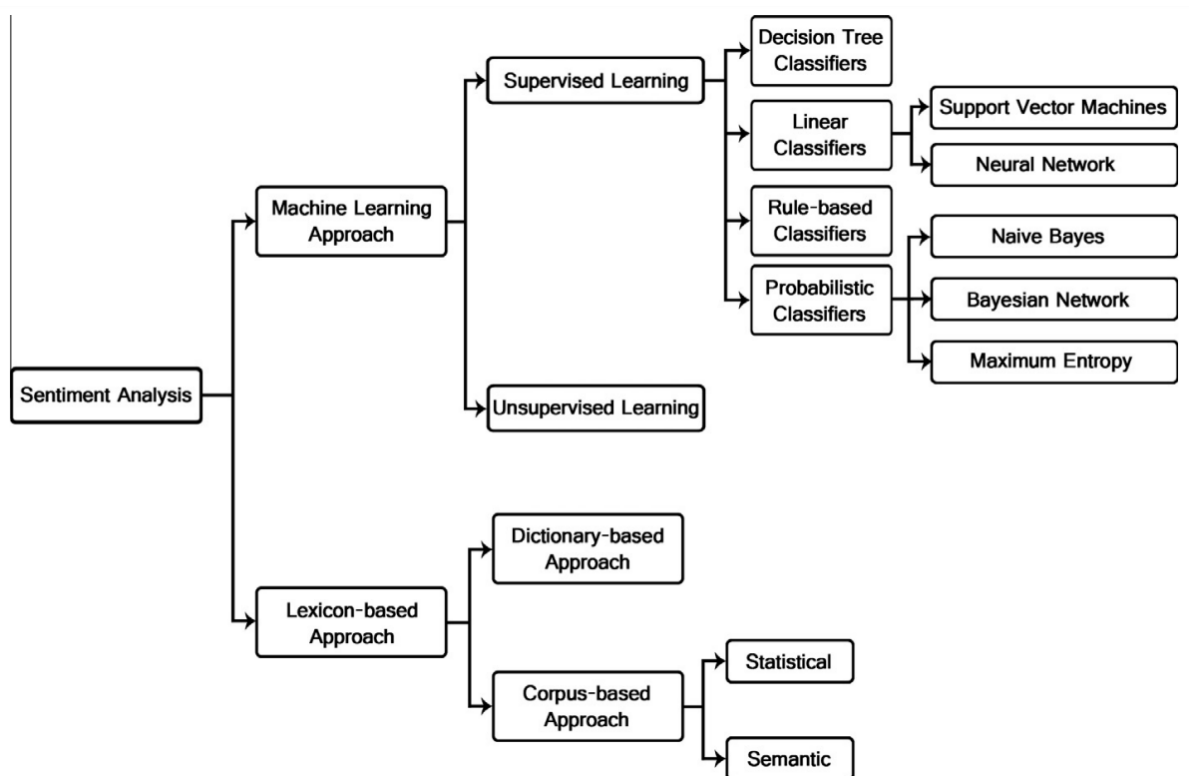
A AS possui três níveis de classificação: nível de documento, nível de sentença e baseada em aspecto. No nível de documento, o documento inteiro é considerado como uma unidade básica de informação, no nível de sentença, o objetivo é identificar o sentimento em cada sentença do documento. Já o nível do aspecto se concentra em identificar o sentimento relacionado a determinados aspectos do texto (MEDHAT; HASSAN; KORASHY, 2014). Neste trabalho, será considerado apenas o nível do documento para as AS, sendo que, no caso deste estudo, um documento é um *tweet*.

A Figura 4 mostra as técnicas existentes para AS. O processo de desenvolvimento de sistemas de AS normalmente segue um conjunto de etapas, embora as etapas implementadas podem variar dependendo da tarefa em questão e da qualidade dos dados. O pipeline típico para o desenvolvimento de algoritmos de AS pode ser dividido nas seguintes etapas:

1. **Coleta de dados:** a primeira etapa é coletar dados de texto. Esses dados podem ser dos mais variados tipos, no entanto, é comum em tarefas de AS a utilização de *reviews* de sites e de postagens em mídias sociais, como o Facebook e o Twitter.
2. **Pré-processamento de texto:** como visto na Seção 2.3.1, os dados de texto geralmente contém ruídos e devem ser pré-processados para transformar os dados em um formato propício para os algoritmos de AS.
3. **Extração de *features*:** os dados de texto são transformados em recursos numéricos que podem ser usados como entrada para um modelo de AM ou baseado em regras. Técnicas comuns para extração de *features* incluem *bag-of-words*, *n-grams* e *word embeddings*.

4. **Treinamento de modelo/Definição das regras:** esta etapa é diferente dependendo do tipo de algoritmo que está sendo utilizado:
 - Baseado em AM: será realizado o treinamento do modelo de AM.
 - Baseado em regras: definição e implementação do conjunto de regras a ser utilizado.
5. **Avaliação do modelo:** o modelo é avaliado em um conjunto de dados de teste separado com o objetivo de obter métricas que podem ser utilizadas para comparar o desempenho do modelo desenvolvido em relação a outros modelos pré-existentes.

Figura 4 – Técnicas de classificação de sentimentos



Fonte: (MEDHAT; HASSAN; KORASHY, 2014)

3.1 Classificação de Emoções

A Classificação de Emoções trata do reconhecimento e categorização de emoções, um aspecto crucial no âmbito da AS. O estudo de emoções e a sua subsequente classificação auxiliam na compreensão dos sentimentos humanos expressos na linguagem, constituindo-se, assim, uma ferramenta vital para a AS e para a interpretação contextual do discurso. Nesta seção, serão abordadas duas metodologias para a classificação de emoções: o Mapeamento de Emoções em Espaços Afetivos (Seção 3.1.1) e a Taxonomia de Emoções (Seção 3.1.2).

3.1.1 Mapeamento de emoções em Espaços Afetivos

O mapeamento de emoções em Espaços Afetivos é uma abordagem multidimensional para representar e entender emoções. Nestes modelos, as emoções são posicionadas em um “espaço” cujos eixos representam dimensões ou domínios emocionais fundamentais. Essas representações espaciais permitem visualizar as emoções de uma maneira que reflete a forma como elas são experienciadas e expressas pelos humanos, fornecendo um meio intuitivo de explorar a complexidade e a inter-relação das emoções.

Existe um grande esforço dedicado à investigação da dimensionalidade do domínio da emoção. Os modelos mais conhecidos consistem em usar um espaço geométrico bidimensional para o mapeamento das emoções, refletindo os eixos da valência e excitação (RUSSELL, 1980; RUSSELL; LEWICKA; NIIT, 1989). No entanto, alguns estudos também identificaram uma terceira dimensão adicional, mais frequentemente rotulada como poder ou controle (FONTAINE et al., 2010; SHAVER et al., 1987; SHAVER; WU; SCHWARTZ, 1992). De forma ainda mais contemporânea, pesquisas têm apontado para a descoberta de um quarto eixo emocional, denominado surpresa (TRNKA et al., 2021), adicionando ainda mais complexidade à representação dimensional das emoções.

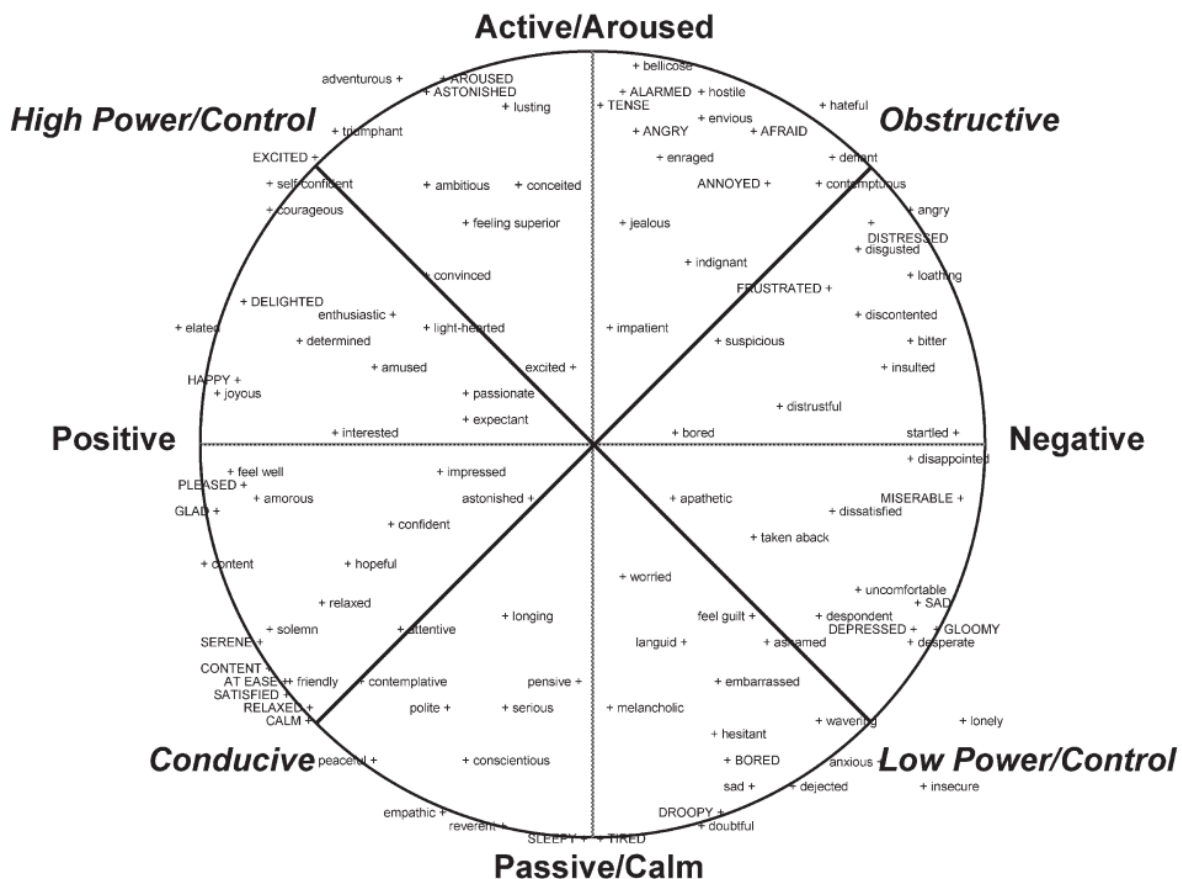
A Figura 5 mostra o Espaço Semântico de Scherer, um exemplo do mapeamento de emoções no espaço afetivo (SCHERER, 2005). O diagrama retrata o mapeamento dos marcadores utilizados por Russell em 1983 (RUSSELL, 1983) em sua proposta de representar emoções em um espaço circunplexo de valência por atividade/excitação (termos em letras maiúsculas). Adicionado a essa representação, encontra-se sobreposta a estrutura bidimensional (eixo de poder/controle por obstrução), derivada de classificações de similaridade de 80 termos emocionais alemães (símbolos “+”, termos em minúsculas, traduzidos para o inglês), oriunda de um estudo anterior (SCHERER, 1984) que corroborou a suposição de que o espaço semântico pode ser organizado por critérios de avaliação. Os símbolos “+” indicam a localização exata dos termos no espaço bidimensional.

A seguir será fornecida uma explicação detalhada para cada eixo mostrado na Figura 5.

- **Eixo de Valência (Positivo - Negativo):** Este eixo representa o quão positiva ou negativa é uma emoção. Por exemplo, a felicidade é colocada na extremidade positiva do eixo, enquanto a tristeza é colocada na extremidade negativa;
- **Eixo de Ativação (Ativo/Estimulado - Passivo/Calmo):** Este eixo representa o nível de energia ou intensidade da emoção. Por exemplo, emoções como a raiva ou o entusiasmo, que são altamente energéticas e estimulantes, são colocadas na extremidade ativa do eixo. Em contrapartida, emoções como a calma ou a apatia, que são de baixa energia ou relaxantes, são colocadas na extremidade passiva do eixo;

- **Eixo de Obstrução (Obstrutivo - Condutivo):** Este eixo está relacionado à avaliação da emoção em termos de sua “condução” ou “obstrução” para atingir os objetivos. Emoções “conducentes” seriam aquelas que promovem ou facilitam a consecução de um objetivo, enquanto emoções “obstrutivas” seriam aquelas que interferem ou impedem a consecução de um objetivo. Por exemplo, o medo é colocado na extremidade do eixo obstrutivo, enquanto a calma é colocada na extremidade condutiva do eixo;
- **Eixo de Controle (Alto Poder/Controle - Baixo Poder/Controle):** Este está relacionado à percepção de controle ou poder que a emoção confere a um indivíduo. Emoções que conferem uma sensação de alto poder ou controle seriam aquelas que fazem um indivíduo se sentir no comando ou dominante. Ao contrário, emoções que conferem uma sensação de baixo poder ou controle seriam aquelas que fazem um indivíduo se sentir impotente ou submisso. Por exemplo, emoções como coragem e confiança são consideradas de alto controle, enquanto emoções como ansiedade e insegurança, que usualmente deixam o indivíduo com a sensação de impotência frente a uma situação são consideradas de baixo controle.

Figura 5 – Espaço Semântico de Scherer



Fonte: (SCHERER, 2005)

3.1.2 Taxonomia de emoções

Taxonomias de emoções são sistemas que organizam e categorizam as emoções com base em suas características, como seus processos psicológicos subjacentes, expressões comportamentais e experiências subjetivas. Elas fornecem um conjunto padronizado de rótulos ou termos para diferentes estados emocionais, permitindo uma identificação e medição mais fáceis de emoções em vários contextos, incluindo pesquisa psicológica, interações sociais e modelos de PLN. As taxonomias de emoções podem variar em relação ao número e aos tipos de emoções incluídas.

Ao utilizar uma taxonomia de emoções, os algoritmos de AS podem identificar emoções específicas presentes no texto, como felicidade, raiva, medo ou tristeza, ao invés de apenas a polaridade positiva/negativa. Essa abordagem fornece informações mais detalhadas sobre as atitudes e opiniões do autor do texto.

No entanto, vale ressaltar que todas as taxonomias de emoções carregam intrinsecamente vieses que refletem a cultura, o contexto e a linguagem em que foi criada e treinada. Cada taxonomia é um produto do seu ambiente, condicionada pelos entendimentos e experiências emocionais que são mais prevalentes ou consideradas mais relevantes naquela cultura específica. Um exemplo ilustrativo é a emoção “saudade”, que está presente na Língua Portuguesa, no entanto, não existe um termo em Inglês que represente esse estado emocional. Assim, se uma taxonomia de emoções fosse criada por nativos de língua inglesa, é pouco provável que o termo “saudade” fosse incluído. Isso ilustra como as taxonomias de emoções são limitadas não apenas pelas palavras disponíveis em um idioma, mas também pelos estados emocionais que uma determinada cultura valoriza ou reconhece.

Esta seção abordará alguns exemplos de taxonomias de emoções comumente usadas, incluindo o modelo Emoções Básicas de Paul Ekman (seção 3.1.2.1) e a taxonomia desenvolvida para o GoEmotions (seção 3.1.2.2).

3.1.2.1 Modelo de Emoções Básicas de Paul Ekman

O modelo de emoções básicas de Paul Ekman é uma teoria bem conhecida das emoções que postula a existência de seis emoções básicas que são universalmente reconhecidas e expressas por pessoas em todas as culturas. Essas emoções são felicidade, tristeza, raiva, medo, surpresa e nojo. Segundo Ekman, essas emoções são inatas e biologicamente determinadas, pois são reconhecidas pelas pessoas independentemente de sua cultura ou idioma (EKMAN, 1992).

A teoria de Ekman foi desenvolvida por meio de uma série de estudos que envolviam mostrar aos participantes fotografias de expressões faciais e pedir-lhes que identificassem as emoções que viam. Sua pesquisa mostrou que essas seis emoções foram consistentemente identificadas em todas as culturas e exibidas por meio de expressões faciais semelhantes.

Com o tempo, Ekman expandiu sua teoria para incluir desprezo, o que elevou o total para sete expressões faciais de emoções universais.

Uma das grandes vantagens do Modelo de Emoções Básicas desenvolvido por Ekman é que ele propõe emoções que são universalmente reconhecidas e expressas de maneira semelhante em diversas as culturas, além ser um modelo simples e claro de entender. No entanto, devido a sua simplicidade, este modelo não contempla toda a gama de emoções experienciadas pelos seres humanos, podendo simplificar demais a natureza complexa da experiência e expressão emocional.

3.1.2.2 GoEmotions

O GoEmotions¹ (DEMSZKY et al., 2020) é o maior dataset manualmente anotado do mundo e um modelo BERT pré-treinado para Análise de Sentimentos desenvolvido pela *Google Research*². Ele possui 54 mil comentários do Reddit³ cuidadosamente selecionados e rotulados por humanos para 27 categorias de emoções ou neutro. Este dataset foi usado para realizar o ajuste fino de um modelo BERT, resultando em um modelo capaz de identificar as emoções presentes no texto.

Para a construção do corpus, foi utilizado um conjunto de dados resultante do *reddit-data-tools project*⁴. A partir dele, foram selecionados os comentários de *sub-reddits* com pelo menos 10 mil membros. Por conta do Reddit ser uma plataforma formada majoritariamente por jovens do sexo masculino, ela tem tendência em possuir comentários com linguagem tóxica e violenta. Para obter um corpus mais generalizável e que não reforce vícios de linguagem e de comportamentos demográficos, a etapa de curadoria e limpeza do corpus consistiu, entre outros, na remoção de comentários nocivos, com termos religiosos, ofensivos e vulgares.

As emoções que compõem a taxonomia do GoEmotions foram escolhidas através de um processo *multiround*, onde em cada rodada uma pequena parte dos dados era anotada e os anotadores poderiam sugerir novos rótulos de emoções em cima de um conjunto de emoções pré-existente. Na primeira rodada, foram utilizadas as emoções identificadas como relevantes por (COWEN; KELTNER, 2017)⁵ para compor o conjunto base de emoções, garantindo que o conjunto incluísse as emoções básicas identificadas por (EKMAN, 1992). Foram anotados cerca de 3 mil exemplos nesta rodada e a taxonomia foi atualizada com base nos resultados. Foram removidas as emoções que eram selecionadas com pouca

¹ <<https://github.com/google-research/google-research/tree/master/goemotions>>

² <<https://research.google/>>

³ <<https://www.reddit.com/>>

⁴ <<https://github.com/dewarim/data-tools-for-reddit>>

⁵ Do original em Inglês: *Admiration, Adoration, Aesthetic Appreciation, Amusement, Anger, Anxiety, Awe, Awkwardness, Boredom, Calmness, Confusion, Craving, Disgust, Empathic Pain, Entrancement, Excitement, Fear, Horror, Interest, Joy, Nostalgia, Relief, Romance, Sadness, Satisfaction, Sexual Desire, Surprise.*

frequência ou que tinham baixa concordância entre os anotadores e adicionadas as emoções frequentemente sugeridas. Este processo foi repetido novamente na segunda rodada, desta vez com 2 mil exemplos anotados. O conjunto final contém 27 emoções mais a categoria neutro.

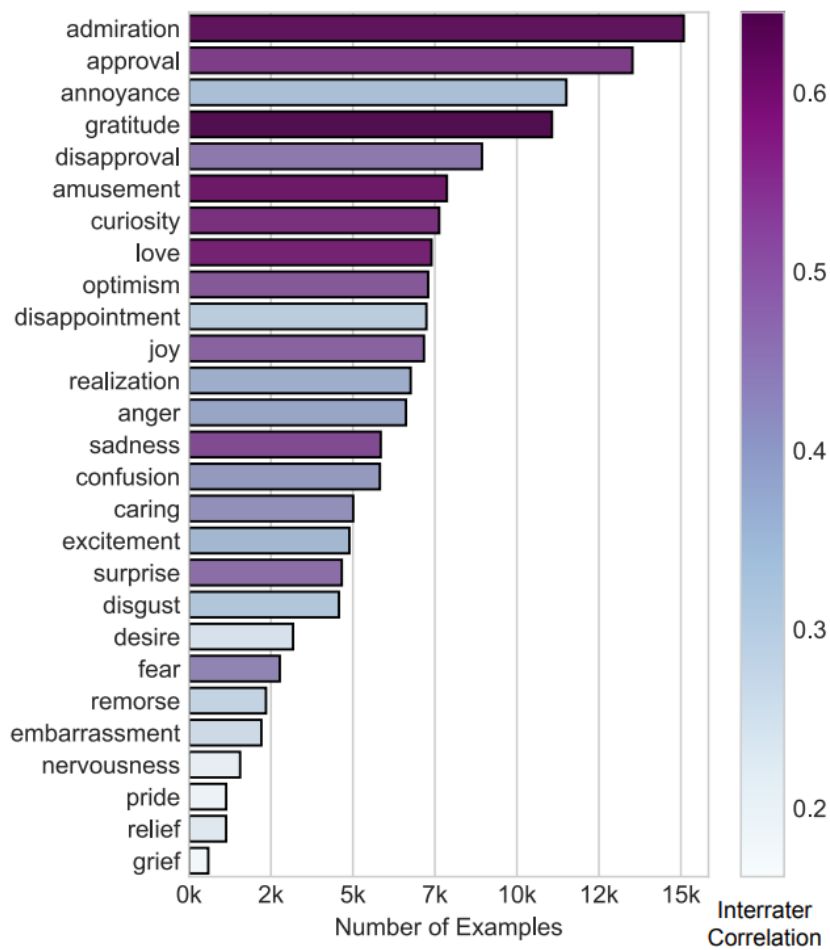
Esta taxonomia foi utilizada como base para anotação do corpus coletado. Todos os comentários foram anotados por 3 anotadores e, caso não houvesse concordância em pelo menos um rótulo de emoção, mais dois anotadores eram designados para a tarefa. Um comentário poderia estar associado a mais de uma emoção, deste que ela estivesse evidentemente expressa no texto. Os comentários muito difíceis de avaliar foram descartados. Ao final, obtve-se um dataset com 54 mil comentários do Reddit rotulados com base na taxonomia do GoEmotions e com concordância de pelo menos dois anotadores.

A Figura 6 mostra as emoções da taxonomia do GoEmotions, ordenadas pelo número de exemplos onde pelo menos um dos três anotadores associaram a emoção à um comentário. É evidente que existe um forte desbalanceamento em relação à frequência das emoções, por exemplo, admiração (*admiration*) é 30 vezes mais frequente que luto (*grief*). No entanto, não foi adotada nenhuma medida para corrigir o desbalanceamento dos dados antes de passar para a etapa de treinamento, e o dataset gerado foi utilizado para realizar o ajuste fino em um modelo *BERT-base*.

3.2 Detecção de emoções em tweets no contexto da política brasileira

A identificação de emoções em publicações do Twitter, particularmente aquelas relativas ao cenário político brasileiro, representa uma área de estudo de importância ascendente. Com a consolidação das mídias sociais – destacando-se o Twitter – como meios de expressão política relevantes, torna-se evidente a necessidade de compreender melhor as manifestações emocionais presentes nesse meio. Graças à sua natureza dinâmica e ao volume substancial de informações geradas, a plataforma do Twitter é uma fonte prolífica de dados para análise de sentimentos e detecção de emoções, propiciando um entendimento mais aprofundado da percepção pública acerca dos atores políticos e de suas decisões.

A relevância deste campo de estudo é ainda mais acentuada no Brasil, um país onde as discussões políticas são frequentemente permeadas por fortes expressões emocionais e onde a população frequentemente recorre às mídias sociais para expressar suas opiniões e sentimentos. A capacidade de capturar e analisar essas emoções pode fornecer uma melhor compreensão sobre o clima político do país. Assim, esta seção visa discutir alguns trabalhos que exploram a detecção de emoções em publicações do Twitter no contexto político brasileiro.

Figura 6 – Categorias de emoções ordenadas por frequência no dataset

Fonte: (DEMSZKY et al., 2020)

3.2.1 Detecção de crenças de grupos

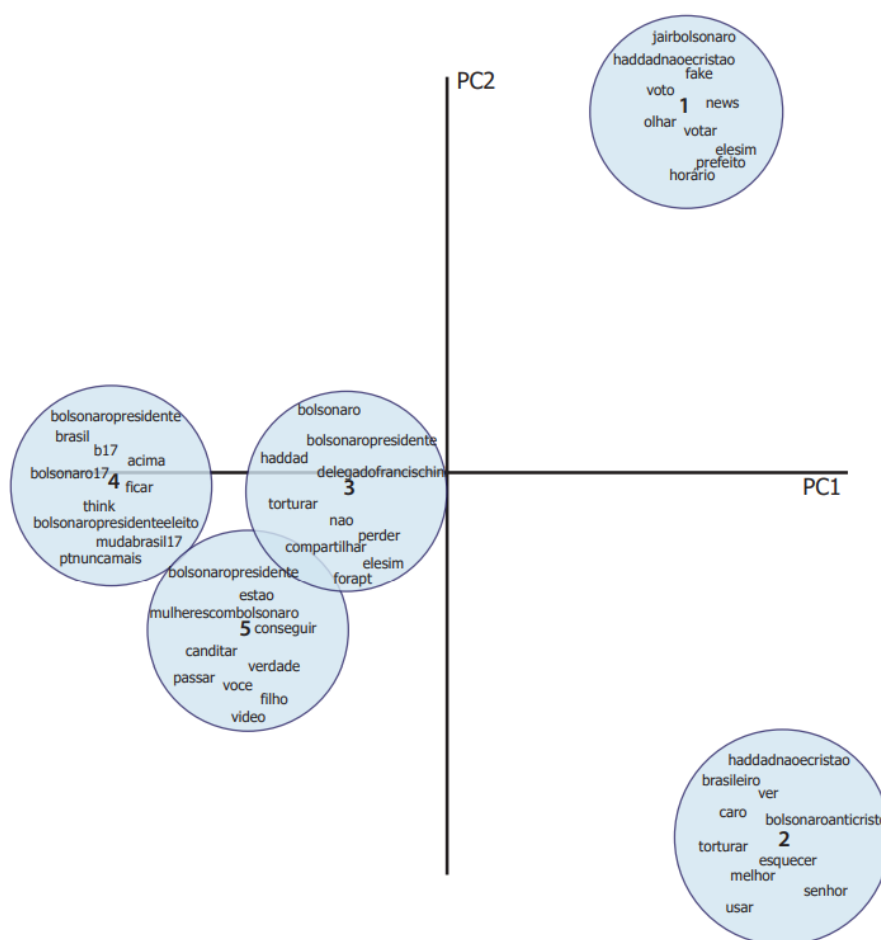
Em (SANTANA; VANIN, 2020) os autores analisaram como os discursos produzidos no Twitter durante a segunda rodada das eleições presidenciais brasileiras de 2018 reforçam o engajamento de grupo com as crenças pessoais. Para isso, foram coletados cerca de 50 mil *tweets* por dia na última semana da campanha eleitoral, e os autores usaram técnicas de modelagem de tópicos e mineração de opiniões para analisar discursos políticos produzidos no Twitter.

O corpus resultante foi submetido a uma etapa de pré-processamento que incluiu normalização, tokenização, remoção de *stopwords*, dentre outros. As palavras foram, então, padronizadas através da lematização, utilizando bibliotecas de processamento de texto, e a abordagem *Word2Vec* (CHURCH, 2017) foi empregada para transformar cada *token* em um vetor numérico que o representa semanticamente. O treinamento do modelo de dados utilizou o modelo *skip-gram* (KRISHNA; SHARADA, 2020). Usando uma ferramenta multiuso para análise de corpus, identificaram-se várias repetições de posts feitos pelos mesmos usuários, levando à remoção de repetições de *tweets* para uma análise mais precisa.

Isso foi considerado um possível indicativo de discursos disseminados e corroborados por bots do Twitter.

Este estudo empregou modelagem de tópicos, utilizando a Análise de Alocação Dirichlet Latente (LDA) (BLEI; NG; JORDAN, 2003), permitindo que cada documento fosse interpretado como uma mistura de vários tópicos. Os autores também reconheceram que as polaridades sentimentais dependem dos tópicos ou domínios, e realizaram uma análise focada em discursos politicamente motivados no Twitter. Ao total, foram identificados 5 tópicos, com uma pontuação de coerência de 0.58, de acordo com a medida c_v , uma medida que captura o número ótimo de tópicos. Além disso, foi utilizado um mapa de distância inter-tópicos, mostrado na Figura 7, via escalonamento multidimensional para visualizar a distribuição de tópicos principais.

Figura 7 – Distância Inter-tópicos



Fonte: (SANTANA; VANIN, 2020)

Por fim, este estudo utilizou o SentiLex-PT (CARVALHO; SILVA, 2015), um léxico de sentimentos, para analisar as polaridades expressas nos *tweets* coletados. Extraíram-se opiniões dos textos observados, identificando expressões e palavras classificadas no léxico como positivas ou negativas. Empregando a abordagem Word2Vec, associou-se polaridades a cada tópico, analisando as palavras principais dos tópicos identificados e, cruzando os

dados extraídos do léxico, calculou-se a similaridade do cosseno. Observou-se como os sentimentos interagiam nos tópicos modelados, usando apenas expressões classificadas como positivas ou negativas. Palavras frequentes não-neutras no conjunto observado foram identificadas e analisadas em relação aos grupos dos tópicos. Verificou-se que sentimentos de polaridade negativa têm maior similaridade com referências de oposição, enquanto a polaridade positiva se refere à paridade ideológica das *hashtags*.

Esses achados indicam a existência de um apelo emocional nos discursos digitais de motivação política, ressaltando a cultura de indiferença e confirmando tendências de viés, o que pode levar a uma maior propensão para lembrar, buscar informações ou interpretar fatos de maneiras que confirmem crenças pré-estabelecidas.

3.2.2 Detecção de polaridade no processo de *Impeachment*

Em (SOUZA et al., 2016) os autores fizeram uma comparação entre as técnicas de aprendizado de máquina para inferir a polaridade dos brasileiros em relação ao processo de impeachment da ex-presidente do Brasil, Dilma Vana Rousseff, usando *Support Vector Machine* (SVM) (CORTES; VAPNIK, 1995), *Multinomial Naive Bayes* (MNB) (KIBRIYA et al., 2005) e *Max Entropy* (ME) (COHEN et al., 2019) em uma base de comentários do Twitter. As principais contribuições do trabalho são a avaliação de técnicas de aprendizado de máquina mais adequadas para o cenário e um estudo de caso sobre o processo de impeachment da ex-presidente do Brasil.

Para formar a base de dados, aproximadamente 10 milhões de *tweets* relacionados ao processo de impeachment da ex-presidente do Brasil foram coletados. Para a base de dados de treinamento e validação, uma amostra de 3.167 *tweets* foi selecionada, e os *tweets* foram rotulados por voluntários como contra o impeachment, a favor do impeachment ou indefinido. Informações adicionais, como a data e o usuário que compartilhou o *tweet*, foram disponibilizadas para ajudar no processo de rotulação. A distribuição final dos rótulos foi de 47,80% contra o impeachment, 47,80% a favor do impeachment, e 4,38% indefinidos. O estudo considerou polaridades contra e a favor do governo, mas também incluiu comentários indefinidos.

Após a criação da base de dados, foi realizado o processo de pré-processamento que incluiu eliminação de ruídos, *stopwords*, URLs, menções e *hashtags*. Para o processo de treinamento e teste, efetuou-se a validação cruzada de 10 partes. As métricas usadas no processo de avaliação no desempenho dos modelos foram precisão, revocação (*recall*), F1 e acurácia. A tabela contendo os resultados dos modelos, extraída do artigo, é ilustrada na Figura 8.

O SVM apresentou um desempenho levemente superior quando comparado aos outros modelos. No entanto, todos os classificadores mostraram um bom desempenho, com

F1 acima de 80% para todas as classes. A classe de rótulos “indefinidos” apresentou menor desempenho, dada a sua menor presença no conjunto de dados. A análise concluiu que a abordagem de aprendizado de máquina supervisionada é eficaz para inferir a polaridade em textos em português relacionados a questões políticas.

Figura 8 – Resultado dos modelos avaliados

| | Método | A favor | Contra | Indefinido |
|------------------|--------|---------|--------|------------|
| Precisão | SVM | 0,9571 | 0,9156 | 1,0000 |
| | MNB | 0,8598 | 0,9084 | 1,0000 |
| | ME | 0,9347 | 0,8922 | 0,9166 |
| Revocação | SVM | 0,9178 | 0,9743 | 0,7333 |
| | MNB | 0,9246 | 0,8910 | 0,4666 |
| | ME | 0,8833 | 0,9551 | 0,7333 |
| F1 | SVM | 0,9370 | 0,9440 | 0,8461 |
| | MNB | 0,8910 | 0,8996 | 0,6363 |
| | ME | 0,9084 | 0,9226 | 0,8148 |

Fonte: (SOUZA et al., 2016)

3.2.3 Análise de Sentimentos nas Eleições de 2018

Em (CRISTIANI; LIEIRA; CAMARGO, 2020) os autores aplicaram técnicas de AS para investigar uma possível relação entre a opinião dos usuários do Twitter e os resultados das eleições presidenciais brasileiras de 2018. Os algoritmos de aprendizado de máquina utilizados para a criação dos modelos de classificação foram o Naive Bayes (WEBB, 2010) e SVM (CORTES; VAPNIK, 1995). Posteriormente, os usuários foram agrupados e os resultados comparados com pesquisas e resultados eleitorais. Os dados do Twitter foram coletados durante eventos-chave da campanha eleitoral (debates, entrevistas e os domingos eleitorais), resultando na coleta de mais de 900 mil *tweets* relacionados à eleição.

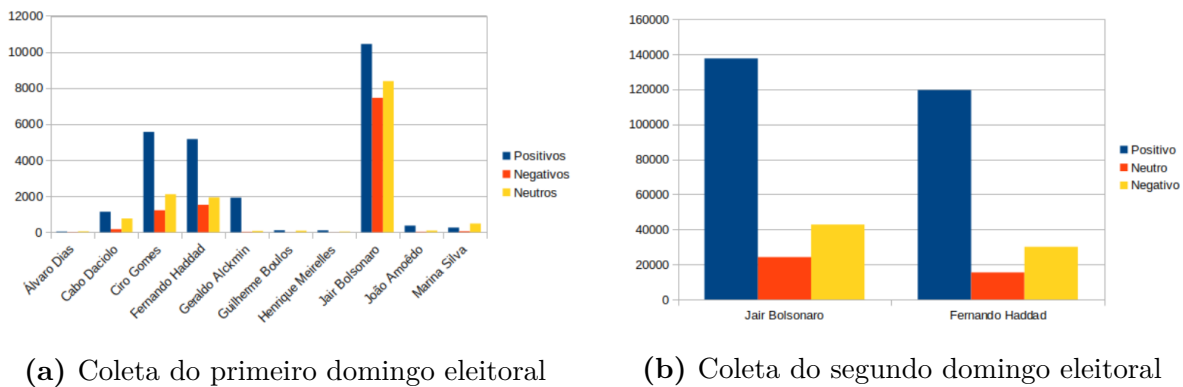
Após a coleta de *tweets* foi realizada uma anotação manual para produzir um conjunto de dados de treinamento e teste. Os *tweets* foram selecionados aleatoriamente e classificados em três polaridades: positivo, negativo e neutro, dependendo do seu conteúdo em relação aos candidatos. Todos os 600 *tweets* anotados foram associados a um candidato específico, utilizando um dicionário de palavras chave. Seguindo a anotação, os dados passaram por uma etapa de pré-processamento para normalizar e preparar os dados, envolvendo etapas de tokenização, remoção de *stopwords*, lematização e TF-IDF.

Em seguida foi realizado um experimento para avaliar o desempenho de dois classificadores, com o intuito de selecionar o melhor para classificar o conjunto de dados completo. A acurácia, precisão, recall e f-measure foram usadas como métricas de avaliação. 70% dos dados foram separados para treinamento e 30% para avaliação. O classificador SVM obteve resultados superiores ao Naive Bayes em todas as medidas avaliadas, obtendo f-measure de 66,6%. O modelo de classificação foi, então, gerado com base no classificador

SVM e o conjunto total de dados foi rotulado, sendo dividido de acordo com os candidatos. Os dados do primeiro domingo eleitoral foram divididos em dez candidatos e os do segundo domingo em dois candidatos, resultando em doze conjuntos de dados classificados pelo modelo.

A Figura 9, extraída do artigo, demonstra a distribuição dos *tweets* e os sentimentos expressos para cada um dos candidatos, levando em consideração as três polaridades analisadas. No primeiro turno das eleições, 48.985 *tweets* referentes aos dez candidatos foram analisados, com cinco deles se destacando em quantidade de mensagens positivas e negativas. No segundo turno, 369.800 *tweets* referentes aos dois candidatos restantes foram analisados, com uma quantidade maior de citações positivas para o candidato Jair Bolsonaro, apesar deste também ter obtido um maior número de citações negativas.

Figura 9 – Resultados obtidos para as coletas referentes ao primeiro e segundo domingo eleitoral.



Fonte: (CRISTIANI; LIEIRA; CAMARGO, 2020)

Por fim, os resultados obtidos do Twitter foram comparados com pesquisas conduzidas por dois institutos de pesquisa durante as eleições presidenciais, o Datafolha⁶ e o Ibope⁷. A comparação mostrou uma correlação entre os candidatos que obtiveram o maior número de *tweets* positivos e aqueles com a maior parte dos eleitores favoráveis, conforme apontado pelas pesquisas. Esta correlação foi observada tanto para os quatro candidatos mais populares no primeiro turno quanto para os dois candidatos no segundo turno. Os dados sugerem que o Twitter é uma fonte valiosa de informação para pesquisas de opinião, com a quantidade de mensagens positivas refletindo os resultados das pesquisas e o resultado final das eleições de 2018.

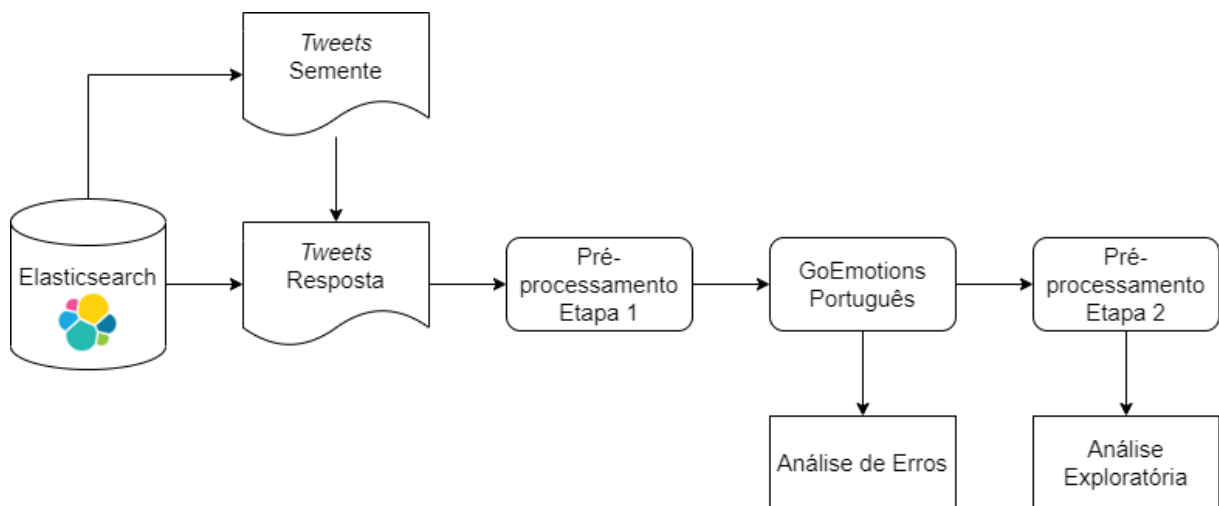
⁶ <<https://datafolha.folha.uol.com.br/>>

⁷ <<https://kantariopemedia.com/brazil/>>

4 Materiais e Métodos

Neste Capítulo são descritos os materiais e os métodos utilizados para o desenvolvimento deste trabalho. A Figura 10 mostra o pipeline principal utilizado nesta pesquisa. Este projeto foi dividido em três módulos principais: extração dos *tweets*, pré-processamento dos dados e inferência de emoções dos *tweets* pelo GoEmotions para Português. A Seção 4.1 apresenta as ferramentas de programação utilizadas para o desenvolvimento do projeto. A Seção 4.2 apresenta o conjunto de dados utilizado, assim como o processo de extração dos dados do *Twitter*. Em seguida, a Seção 4.3 descreve as duas etapas de pré-processamento realizadas e, por fim, a Seção 4.4 apresenta o modelo GoEmotions para Português, utilizado para inferir os sentimentos nos dados coletados.

Figura 10 – Pipeline do projeto



Fonte: Próprio da Autora

4.1 Google Colab e Bibliotecas Python

O Google Colab¹ é plataforma de computação em nuvem, disponibilizada pela Google, que oferece um ambiente de desenvolvimento gratuito para a execução de notebooks Jupyter. Ele permite a colaboração em tempo real e oferece acesso gratuito a GPUs (Unidades de Processamento Gráfico) e TPUs (Unidades de Processamento Tensorial). Esses recursos de hardware especializados são essenciais para a computação paralela e são altamente eficazes em tarefas que envolvem grandes volumes de dados, como o treinamento de modelos de aprendizado profundo.

¹ <<https://colab.research.google.com/>>

No contexto deste projeto, a utilização do Google Colab mostrou-se fundamental, especialmente pela disponibilidade da GPU oferecida pela plataforma. O modelo GoEmotions, utilizado neste projeto, possui complexidades computacionais, que tornam o processo de inferência custoso em relação ao tempo, e a necessidade de recursos de computação de alta qualidade é um pré-requisito para que o modelo fosse viável de ser executado.

Em conjunto com o Google Colab, este projeto também se beneficiou do uso de várias bibliotecas Python, cada uma contribuindo com funções e recursos específicos para o desenvolvimento e implementação da pesquisa.

- **Pandas:** utilizada para manipular os corpus coletados. Esta biblioteca fornece uma interface intuitiva para a organização de dados, o que facilitou a análise e preparação dos corpus para as etapas subsequentes;
- **Elasticsearch:** empregada para capturar os *tweets* que estavam armazenados no Elasticsearch através de queries;
- **NLTK (Natural Language Toolkit):** essa biblioteca foi aplicada na etapa de pré-processamento dos dados, preparando-os para análises posteriores.
- **Scikit-Learn (sklearn):** essa é uma biblioteca de aprendizado de máquina e auxiliou na geração da lista de palavras mais frequentes através do método *bag-of-words*;
- **WordCloud:** utilizada para gerar nuvens de palavras, proporcionando uma representação gráfica das palavras mais frequentes e permitindo uma interpretação visual rápida e intuitiva;
- **Seaborn:** empregada na criação de gráficos, permitindo uma representação visual clara e informativa dos resultados da análise.

Vale ressaltar que as etapas de coleta, pré-processamento, inferência pelo modelo GoEmotions e análise dos dados foram realizadas pela aluna, através da geração de scripts e notebooks do Google Colab. Além disso, o uso de ferramentas de PLN (como as bibliotecas NLTK, wordcloud, sklearn, entre outras) desempenharam um papel fundamental na análise exploratória dos dados e na identificação de erros associados ao modelo GoEmotions. O domínio de tais ferramentas e a compreensão profunda dos modelos neurais, muitas vezes referidos como caixas-pretas, são indispensáveis para evitar vieses e prejuízos, garantindo uma interpretação correta dos dados.

Todos os códigos e análises desenvolvidos ao decorrer desta pesquisa, assim como as bases de dados coletadas estão disponíveis em um repositório no Github² do LALIC³ (Laboratório de Linguística e Inteligência Computacional)

² <https://github.com/LALIC-UFSCar/TCC_Fernanda_Malheiros/tree/main>

³ <<https://sites.google.com/site/lalicufscar/>>

4.2 corpus

Para este trabalho, foi utilizado o dataset coletado pelo *GATE Team*⁴ da Universidade de Sheffield, UK. O corpus coletado contém todos os *tweets* publicados pelo ex-presidente do Brasil, Jair Bolsonaro, assim como todas as respostas que ele recebeu no período de 04/01/2020 até 29/05/2022. Além disso, também foram coletados *tweets* que continham as seguintes *hashtags* relacionadas ao cenário político brasileiro e da pandemia do COVID-19:

- #bolsonarotemraza, #chinavirus, #chinesevirus, #corona, #corona19, #corona2019, #coronav, #coronavirus, #coronavirus19, #coronavirus2019, #coronavirusbrasil, #coronavirusnobrasil, #coronaviruspandemia, #coronavirusplanta, #coronavírus, #coronavírusplanta, #covid, #covid-19, #covid19, #covid2019, #covid_19, #filme-seuhospital, #hidroxicloroquina, #sarscov2, #sintomascoronavirus, #sintomascovid, #vachina, #vacina, #viruschines, #wuhanvirus

A aluna obteve acesso a este dataset através da BEPE (Bolsa de Estágio em Pesquisa no Exterior), financiada pela FAPESP⁵ (Fundação de Amparo à Pesquisa do Estado de São Paulo) e realizada da Universidade de Sheffield, sob supervisão da Profa. Dra. Carolina Scarton.

Em Maio de 2022, o *GATE Team* havia coletado aproximadamente 34 milhões de *tweets*, que estavam indexados no Elasticsearch⁶ (GORMLEY; TONG, 2015). O Elasticsearch é uma ferramenta construída com base no Apache Lucene e é um mecanismo distribuído de pesquisa e análise. Com ela, é possível utilizar a API ou ferramentas de ingestão para enviar dados no formato JSON. O Elasticsearch então armazena automaticamente o documento original e adiciona uma referência pesquisável ao documento. Desta forma, utilizando a API do Elasticsearch é possível realizar buscas e recuperar o documento, que no caso desta pesquisa são os *tweets*.

Além disso, em conjunto com o Elasticsearch é possível utilizar o Kibana, uma ferramenta de visualização de dados. A Figura 11 mostra o painel principal do Kibana. Até dezembro de 2022, haviam sido coletados mais de 65 milhões de *tweets*.

4.2.1 Extração dos *tweets*

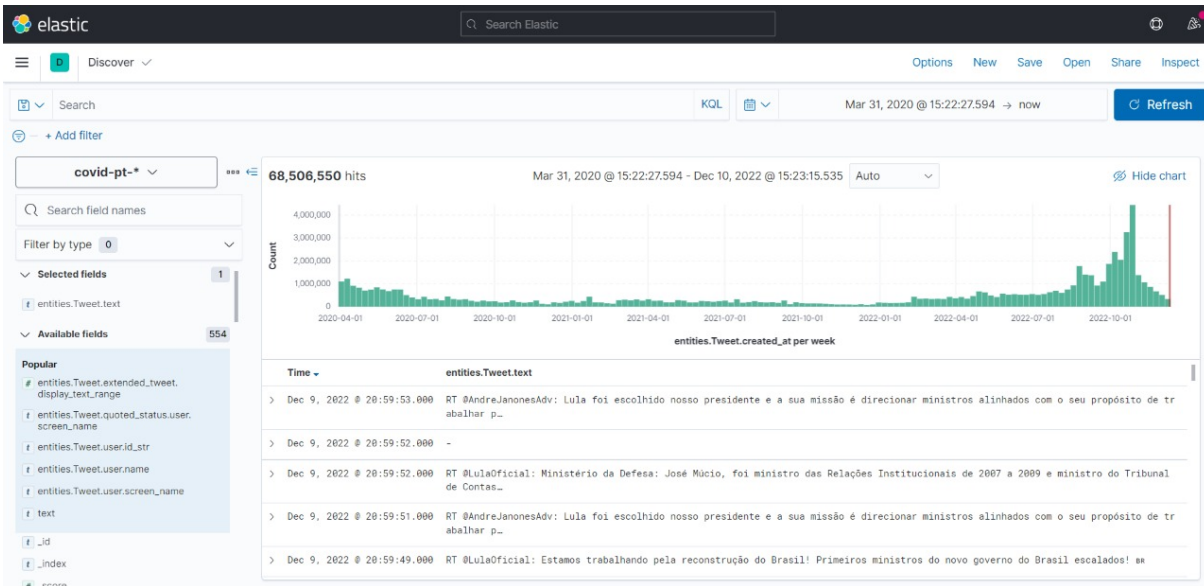
Embora tivessem sido coletados mais de 65 milhões de *tweets*, para esta pesquisa interessa apenas aqueles postados pela conta do Bolsonaro (chamados de *tweets* semente) e

⁴ <<https://gate.ac.uk/>>

⁵ <<https://fapesp.br/>>

⁶ <<https://www.elastic.co/>>

Figura 11 – Painel do Kibana



Fonte: Próprio da Autora

as respostas que ele obteve em cada um. Sendo assim, a API do Elasticsearch foi utilizada para extrair os *tweets* relevantes.

4.2.1.1 Tweets semente

O primeiro passo foi coletar todos os *tweets* semente. Para isso, foi realizada uma pesquisa na base de dados utilizando o *id* da conta do Bolsonaro. Em seguida, todos os *tweets* foram salvos em um arquivo CSV contendo os seguintes campos:

- **tweet_id**: Id do *tweet* semente;
- **user_id**: Id do usuário neste caso é o id da conta do Bolsonaro;
- **created_at**: A data e hora de criação do *tweet*, incluindo o dia da semana, o dia, mês e ano, assim como a hora, minuto e segundo;
- **tweet_text**: O texto do *tweet*;
- **is_retweet**: Um campo booleano que indica se o *tweet* postado era um *retweet*.

Por fim, após realizada a coleta, a API do Google Drive⁷ foi utilizada para fazer o carregamento automático do arquivo gerado para uma pasta especificada do Google Drive. Este arquivo contém 4305 *tweets* semente postados no período de 04/01/2020 até 29/05/2022.

⁷ <<https://developers.google.com/workspace/guides/create-credentials#api-key>>

4.2.1.2 Respostas aos *tweets* semente

O próximo passo foi coletar cada resposta para todos os *tweets* semente (chamados de *tweets* resposta). Primeiramente, foram descartados todos os *retweets*, já que para esta pesquisa apenas interessavam os *tweets* originais. Em seguida, foi dado como entrada ao Elasticsearch o id de cada *tweet* semente. Inicialmente, foram gerados arquivos CSV separados com todas as respostas para cada *tweet* semente, contendo os seguintes campos:

- **tweet_id**: Id do *tweet* resposta;
- **created_at**: A data e hora de criação do *tweet*, incluindo o dia da semana, o dia, mês e ano, assim como a hora, minuto e segundo;
- **tweet_text**: O texto do *tweet*;
- **id_tweet_reply**: Id do *tweet* semente ao qual este *tweet* está respondendo.

Com o uso da API do Google Drive, todos os arquivos gerados foram automaticamente carregados em uma pasta especificada. Por fim, todos os arquivos resultantes foram concatenados em um único arquivo CSV, contendo todas as respostas de todos os *tweets* semente coletados. O arquivo final possui 4.971.545 *tweets* postados no período de 01/04/2020 até 29/05/2022.

4.3 Pré-processamento

Como visto na Seção 2.3.1, a etapa de pré-processamento e limpeza dos dados é responsável pela redução de ruídos de diversos tipos que podem estar presentes no texto. A implementação desta etapa é fundamental para que se obtenha um bom resultado nos modelos de Aprendizado de Máquina (AM) e em diversas outras tarefas de Processamento de Linguagem Natural (PLN), e suas etapas podem variar dependendo do resultado que se deseja obter.

Para este projeto, foram realizadas duas etapas de pré-processamento distintas. A primeira etapa (Etapa 1) tinha o objetivo de preparar o corpus para o modelo pré-treinado do GoEmotions para Português. Assim, este processo foi simples e consistiu nos seguintes processos:

- **Padronização do texto**: todos os caracteres foram transformados para caixa baixa e as linhas e espaços vazios foram eliminados;
- **Remoção de tags e links**: nesta etapa foram eliminados todos os links, marcações de usuários, e textos que começavam com a *string* “RT”.

Já a segunda etapa (Etapa 2) do pré-processamento tinha o objetivo de preparar o corpus para análises linguísticas. Desta forma, foi necessária a realização de um pré-processamento mais minucioso, de forma a eliminar o máximo de ruído possível. Para isso, além dos processos já aplicados na Etapa 1, foram realizadas as seguintes operações:

- **Remoção de *stopwords*:** remoção de *stopwords* a partir da lista pré-definida pelo NLTK⁸ (LOPER; BIRD, 2002)
- **Remoção de ruídos:** remoção de todas as pontuações, caracteres especiais e símbolos, além de caracteres unitários e quebra de linha.

A Tabela 1 apresenta o resultado obtido com as duas etapas de pré-processamento para três exemplos.

Tabela 1 – Resultados das etapas de Pré-processamento

| Texto original | Etapa 1 | Etapa 2 |
|---|--|--|
| @jairbolsonaro Nem Trump acredita mais em Bolsonaro. https://t.co/tEtHyyGbgE https://t.co/T1cA0gYvi4 Questionado sobre atitudes de Bolsonaro, Trump diz que estuda banir voos do Brasil https://t.co/5fBeIV4kW7 https://t.co/yKAajfVFLu | nem trump acredita mais em bolsonaro. questionado sobre atitudes de bolsonaro, trump diz que estuda banir voos do brasil | trump acredita bolsonaro questionado sobre atitudes bolsonaro trump diz estuda banir voos brasil |
| @jairbolsonaro No vídeo, 6:30 - 6:40, fala que vai sancionar os R\$600 para os mais necessitados. Até agora nada. | no vídeo, 6:30 - 6:40, fala que vai sancionar os R\$600 para os mais necessitados. até agora nada. | vídeo fala sancionar necessitados agora nada |
| @jairbolsonaro Estão espalhando que vc entregou sua carta de renúncia. Me diz que é mentira, por favor! | estão espalhando que vc entregou sua carta de renúncia. me diz que é mentira, por favor! | espalhando entregou carta renúncia diz mentira favor |

4.4 GoEmotions para Português

Para o desenvolvimento da versão em Português do GoEmotions (HAMMES; FREITAS, 2021), os autores utilizaram o corpus original do GoEmotions em Inglês, e ferramentas de tradução automática para traduzir o corpus para o Português. Para resolver o problema do desbalanceamento entre as classes, foi utilizado o *Class Balance*

⁸ [urlhttps://www.nltk.org/](https://www.nltk.org/)

Loss (CB) (CUI et al., 2019), explicado brevemente na Seção 2.4. Por fim, utilizando o dataset balanceado, foi realizado o ajuste fino de dois modelos BERT pré-treinados para o Português, o *BERTimbau-base* e o *BERTimbau-large*. Os hiper-parâmetros e a divisão do dataset entre treinamento, validação e teste foram mantidos.

A Figura 12 mostra a tabela extraída de (HAMMES; FREITAS, 2021), com os comparativos dos resultados do modelo treinado tanto para versão em inglês (Seção 3.1.2.2), quanto para as versões em Português. Em geral, todos os modelos apresentados possuem resultados de avaliação semelhantes. No entanto, é possível notar um leve aumento na média macro do F1-score ao comparar o modelo original (*BERT-base*) com o modelo para Português (*BERTimbau-base*), que são modelos treinados com a mesma quantidade de parâmetros. Essa melhora de desempenho pode ser atribuída ao método de balanceamento utilizado na versão em Português.

Os modelos obtiveram melhor desempenho nas emoções com polaridade positiva, como gratidão (*gratitude*), diversão (*amusement*) e amor (*love*) e obtiveram menor pontuação F1 em emoções de frequência mais baixa, como luto (*grief*), alívio (*relief*) e percepção (*realization*). Isso significa que o modelo acaba confundindo as emoções menos frequentes por emoções mais frequentes e relacionadas em sentimento e intensidade, como luto (*grief*) e tristeza (*sadness*). Um outro ponto importante é que, enquanto no modelo em inglês, a emoção luto (*grief*) havia obtido F1-score de 0.0, na versão em português, com os dados balanceados, foi possível obter um F1-score de 0.22 no modelo base e 0.25 no modelo large.

Para este trabalho, foi utilizada a versão *BERTimbau-large* em Português do modelo do GoEmotions para determinar as emoções presentes nos *tweets* resposta. Este modelo possui como saída uma lista com as 27 emoções ou neutro, juntamente com um score para cada emoção. Quanto maior for o score de uma emoção, maior a probabilidade do texto expressar aquela emoção. O dataset pré-processado na primeira etapa (veja Seção 4.3) foi passado como entrada para o modelo do GoEmotions e a emoção com o maior score, assim como seu score, foram salvos no arquivo CSV com todos os *tweets*.

Figura 12 – Resultados do GoEmotions

| 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 | 0,00 | 0,13 | 0,67 | 0,22 | 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 |
| média macro | 0,40 | 0,63 | 0,46 | 0,45 | 0,55 | 0,48 | 0,45 | 0,57 | 0,50 |

Fonte: (HAMMES; FREITAS, 2021)

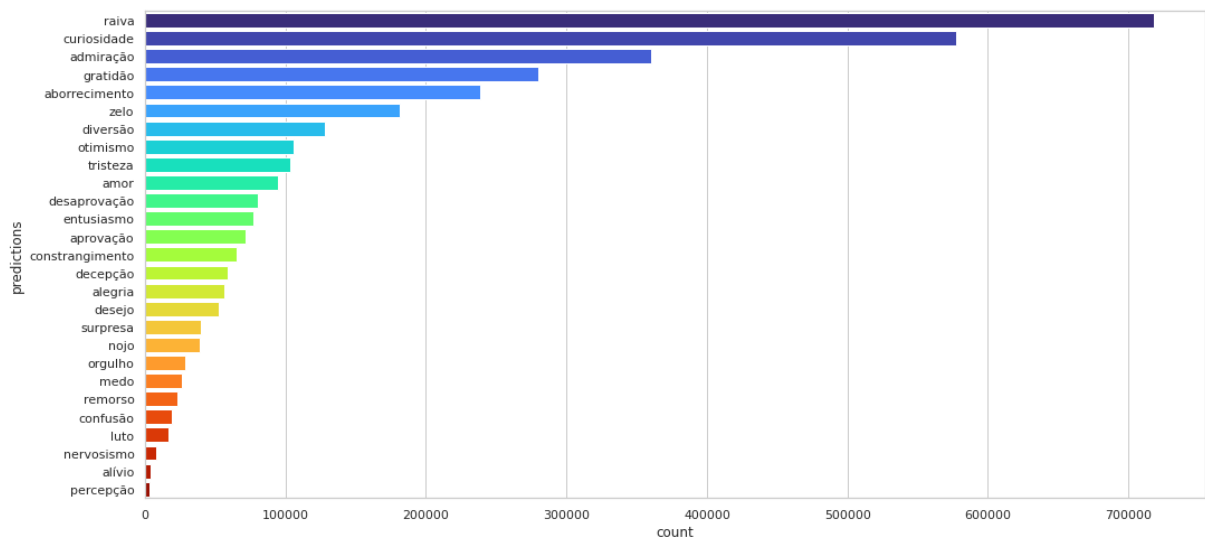
5 Resultados e Discussões

Este Capítulo apresenta os resultados da análise da saída do GoEmotions para Português, nos *tweets* resposta, conforme descrito no Capítulo 4. A análise dos resultados está dividida em duas partes, a primeira apresenta uma análise exploratória dos resultados obtidos (Seção 5.1); já a segunda parte apresenta a análise de erros para as emoções mais frequentes do corpus (Seção 5.2).

5.1 Análise Exploratória

Os resultados apresentados na Figura 13 ilustram a distribuição da quantidade de instâncias para cada emoção no conjunto dos *tweets* resposta, que contém 4.971.545 instâncias. Cada *tweet* foi rotulado com apenas uma emoção, e é evidente na figura que as classes de emoções não estão equilibradas no conjunto de dados. Por exemplo, a emoção mais predominante no corpus (“raiva”) aparece cerca de 247 vezes a mais do que a emoção mais rara (“percepção”). Especificamente, as emoções que mais aparecem nos *tweets* são “raiva” (718.442 instâncias), seguida de “curiosidade” (577.169 instâncias) e “admiração” (360.348 instâncias). Por outro lado, as emoções “percepção” (2900 instâncias), “alívio” (3680 instâncias) e “luto” (16556 instâncias) foram as menos expressas no corpus.

Figura 13 – Quantidade de instâncias por emoção



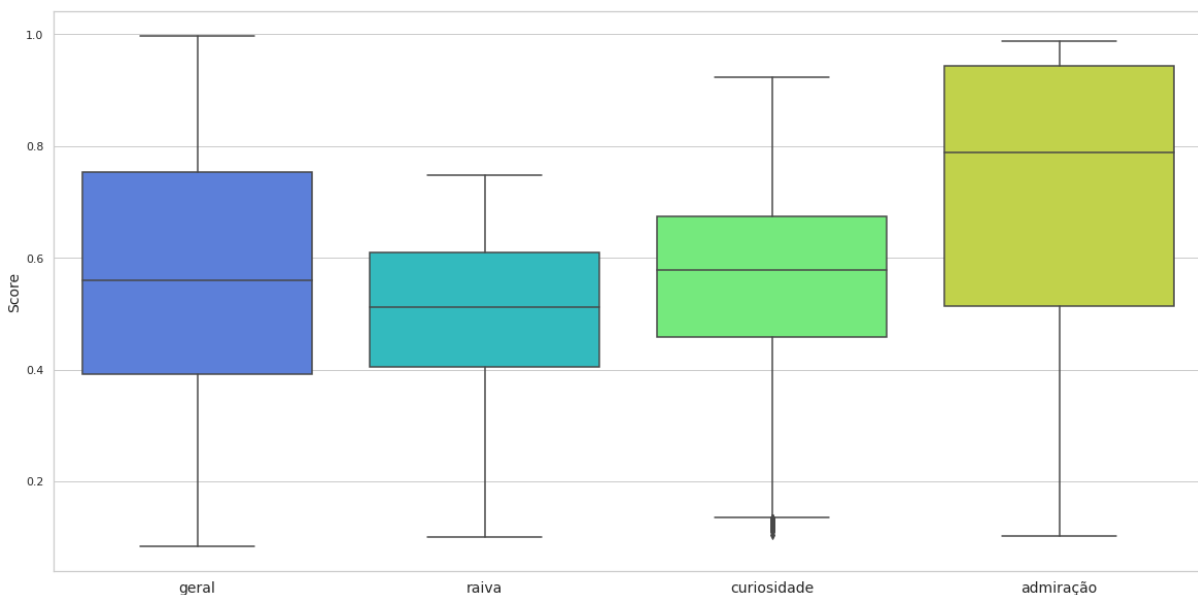
Fonte: Próprio da Autora

Conforme explicado na seção 4.4, cada *tweet* é associado a uma emoção e a um score, que reflete a confiança do modelo em sua classificação. A Figura 14 apresenta quatro

boxplots, que indicam a variação nos scores obtidos para o corpus completo (“geral”) e para as três emoções mais frequentes (“raiva”, “curiosidade” e “admiração”).

Todos os *boxplots* apresentam um valor mínimo aproximado de 0.1, o que sugere que o modelo possui pelo menos um exemplo com baixa confiabilidade para todas as emoções investigadas. Com exceção da emoção “raiva”, cujo valor máximo é ligeiramente superior a 0.7, todas as demais emoções em foco demonstram valores máximos mais elevados, em torno de 0.9. Entretanto, ao considerar as médias, destaca-se a emoção “admiração” com uma média mais elevada, aproximadamente 0.8, enquanto as outras emoções apresentam médias em torno de 0.5-0.6, o que implica que o modelo possui maior confiança ao classificar um *tweet* com a emoção “admiração”. Adicionalmente, a emoção “admiração” também se sobressai pela presença de uma caixa mais alongada, indicando maior variabilidade nos scores em comparação com as demais emoções.

Figura 14 – *Boxplot* dos scores por emoção



Fonte: Próprio da Autora

A Tabela 2 apresenta as palavras mais frequentes para o corpus completo, denominado “Geral”, e para as três emoções mais comuns, nomeadamente “Raiva”, “Curiosidade” e “Admiração”. Observa-se que todas as emoções em questão possuem palavras fortemente relacionadas ao contexto político brasileiro, evidente por palavras como “presidente”, “bolsonaro”, “governo” e “brasil”.

A categoria “Geral” contém termos mais neutros e abrangentes, tais como “presidente”, “Brasil”, “deus” e “povo”, o que é coerente, uma vez que esta categoria engloba todas as emoções do corpus, sem focar um contexto emocional específico.

Notavelmente, a emoção “Raiva” é caracterizada por palavras de natureza emocional altamente negativa, como “merda”, “cu”, “puta” e “vagabundo”, refletindo um sentimento

“raiva” era esperada para ter uma precisão de apenas 0,4 no modelo *BERTimbau-large PT*, substancialmente menor do que a precisão observada.

Em contrapartida, a performance do modelo na identificação da emoção “Curiosidade” exibiu uma queda na precisão, que desceu de uma projeção inicial de 0,45 para um valor efetivo de 0,36. Mesmo que o score médio para os acertos tenha se mostrado marginalmente superior (0,54), o modelo apresentou um número consideravelmente maior de erros (64 erros), com um score médio próximo de 0,51. Estes resultados sugerem uma potencial dificuldade do modelo em distinguir a emoção “Curiosidade” de outras emoções.

Já emoção “Admiração” obteve uma precisão de 0,62, em conjunto com um score médio de 0,74. A precisão efetiva do modelo se mostrou consistente com as expectativas iniciais, estabelecidas em 0,60. Ainda assim, é relevante notar que esta emoção apresentou a maior discrepância entre o score médio para acertos, que foi de 0,74, e o score médio para erros, que atingiu apenas 0,59.

Vale ressaltar que a utilização do modelo GoEmotions para identificar emoções em *tweets* da política brasileira, apresenta algumas limitações importantes. Primeiramente, a taxonomia de emoções do GoEmotions foi desenvolvida com base em dados do Reddit, uma plataforma que tem um contexto sociocultural, demográfico e até linguístico muito diferente do Twitter e, em particular, dos usuários que interagem com posts políticos no Brasil. Além disso, o treinamento do GoEmotions também foi realizado utilizando esses dados anotados do Reddit, tornando o modelo potencialmente enviesado para interpretar emoções de acordo com as normas e padrões predominantes naquela comunidade. Portanto, a generalização do desempenho do modelo para analisar *tweets*, que são breves e muitas vezes mais ambíguos pode representar um desafio para o modelo.

Nas próximas seções, será possível encontrar exemplos de acertos e erros na identificação das principais emoções pelo modelo. Estes exemplos foram selecionados para permitir uma melhor compreensão de como o modelo determina cada emoção e quais podem ser as possíveis causas para suas falhas.

Tabela 3 – Resultado da análise manual de 100 instâncias de cada classe (Raiva, Curiosidade, Admiração)

| Raiva | | Curiosidade | | Admiração | |
|----------------|--------------|----------------|--------------|----------------|--------------|
| Acertos | Score | Acertos | Score | Acertos | Score |
| 82% | 0,52 | 36% | 0,54 | 62% | 0,74 |
| Erros | Score | Erros | Score | Erros | Score |
| 18% | 0,43 | 64% | 0,51 | 38% | 0,59 |

5.2.1 Raiva

A Tabela 4 mostra três exemplos de acertos e erros para a emoção “Raiva”. Os exemplos de acertos revelam que o modelo foi eficaz ao identificar contextos de agressividade verbal e uso de linguagem inflamatória. Tais exemplos possuem em comum o uso de insultos, palavrões e expressões de raiva direta, o que provavelmente auxiliou o modelo a classificá-los corretamente. Essa tendência indica que o GoEmotions parece bem sintonizado com palavras e frases negativas intensas, uma característica útil para a identificação a emoção de raiva.

Por outro lado, os erros podem ser atribuídos a uma variedade de fatores. O primeiro exemplo mostra uma mensagem que, embora pareça expressar uma forma de frustração e desaprovação, não é tão agressiva ou direta quanto os exemplos de acertos. O segundo exemplo apresenta uma combinação de crítica e sarcasmo, que pode ter confundido o modelo. O último exemplo apresenta o palavrão “foda” em um contexto positivo, com uma *hashtag* de aprovação ao final, o que pode ter levado o modelo a não classificar a raiva corretamente.

Em geral, os resultados sugerem que o modelo tem um bom desempenho na identificação da raiva, especialmente quando esta é expressa de forma direta e intensa e com palavras inflamatórias. No entanto, exemplos mais sutis de desaprovação ou mistos de raiva, que muitas vezes não são identificados como tal pelo GoEmotions (FN), representam um desafio maior para o modelo¹.

Tabela 4 – Exemplos de erros e acertos para Raiva

| | Acertos (VP) | Erros (FP) |
|---|--|--|
| 1 | Foda é acordar todo santo dia com vontade de quebrar sua cara seu filho da puta | Presidente, não vá vender o nosso país pra China! Misericórdia! |
| 2 | Vá se foder traidor da pátria seu comunista de armário!!!! Votei em você em co 1 ano de governo conseguiu né fazer mais nojo do que o Luladrão seu desgraçado . Parece que votamos no Haddad seu merda! PGR é STF indicando petistas porra . Filinho com pé na cadeia seu merda . ADELIOO | O pior mentiroso é aquele q acredita nas próprias mentiras. #JAILBolsonaro? Ngm te aguenta mais. Nem teus fiéis, até doidamares já se posicionou a favor do uso de máscara e vacina. Tu tá isolado com meia duzia de fascistas e só. |
| 3 | Vai a merda ditardozinho de Bosta | O homi é foda mesmo BRBRBRBRBR #VaiSerNoPrimeiroTurno |

¹ É importante notar que a identificação de casos específicos de FN requereria uma análise profunda e demorada do conjunto de dados, que vai além do escopo deste trabalho.

5.2.2 Curiosidade

A Tabela 5 mostra três exemplos de acertos e erros para a emoção “Curiosidade”. A identificação correta desta emoção nos exemplos de acerto apresentados é caracterizada pelo uso de perguntas diretas e a solicitação de informações. Nestes casos, o modelo foi capaz de detectar corretamente a intenção de buscar entendimento ou esclarecimentos, o que é intrínseco à emoção de curiosidade.

No entanto, esta emoção também foi marcada por erros significativos. O primeiro exemplo de erro é uma pergunta, que em outras circunstâncias poderia ser classificada como curiosidade, mas neste contexto, ela está carregada de ironia. Este *tweet* é referente a um evento político relevante, no qual Fabrício Queiroz, um ex-assessor do filho de Bolsonaro, foi acusado de depositar R\$89.000 na conta da esposa do ex-presidente Bolsonaro². É importante ressaltar que variações dessa frase apareceram frequentemente no corpus. Modelos de linguagem como o GoEmotions podem lutar para distinguir entre curiosidade genuína e perguntas retóricas ou irônicas, e isso parece ser o caso aqui. No segundo exemplo, a pergunta é mais uma provocação do que uma expressão de curiosidade, o que pode ter confundido o modelo. Já o último exemplo é uma declaração crítica que não expressa curiosidade, mas que pode ter sido mal interpretada pelo modelo devido à sua natureza questionadora, evidenciada pela pergunta retórica ao final.

Os resultados mostram que a emoção “curiosidade” foi um desafio para o modelo, especialmente em casos onde a ironia e o sarcasmo estavam presentes, o que é comum em discussões políticas online. Estas nuances linguísticas são complexas e representam uma área onde o modelo poderia ser aprimorado.

5.2.3 Admiração

A Tabela 6 mostra três exemplos de acertos e erros para a emoção “Admiração”. A análise das previsões do modelo para esta emoção sugere um desempenho razoavelmente satisfatório, embora ainda sujeito a erros. Entre os acertos, o modelo foi capaz de identificar corretamente a emoção de admiração em cada um dos três exemplos apresentados, provavelmente devido à presença de palavras e expressões de forte conotação positiva, como “impressionante”, “parabéns” e “show de bola”.

Em contrapartida, a análise dos erros revela, novamente, dificuldades notáveis na interpretação correta de textos irônicos ou sarcásticos. No primeiro exemplo, a presença da palavra “corajoso” levou a uma interpretação errônea, apesar do contexto geral ser crítico. O segundo exemplo de erro é um caso clássico de ironia, onde um elogio aparente (“Parabéns por fazer o mínimo”) na verdade expressa crítica ou censura. No terceiro exemplo, a

² <https://g1.globo.com/rj/rio-de-janeiro/noticia/2020/08/07/fabricio-queiroz-e-a-esposa-repassaram-r-89-mil-para-michelle-bolsonaro.ghtml>

Tabela 5 – Exemplos de erros e acertos para Curiosidade

| | Acertos (VP) | Erros (FP) |
|---|--|--|
| 1 | São Paulo receberá dessas vacinas? | Presidente porque o Fabricio Queiroz depositou 89 mil na conta da sua esposa? |
| 2 | Presidente, o Brasil está precisando com urgência de fabricar tecnologias simples porém que dariam um up na produção de bens e serviços, ex mini tratores bike elétricas etc, a Embraer tem estrutura para isso, pq não por fábricas desses itens no nordeste?? | Preparado para perder a eleição em 2022 presidente??? |
| 3 | Com todo respeito senhor presidente Bolsonaro e o salário como fica e os aposentados e pensionista ? Precisamos de salário que possamos pagar o nosso aluguel , água luz alimentação pois 1100 com os descontos não nós sobre nem pro aluguel nem todos que trabalham ganhar bem | Enquanto o país está à deriva, o chefe de Estado fica passeando de moto, sem máscara, causando aglomeração. Ainda não ficou claro q vc e seu desgoverno já derreteu? |

presença de um termo geralmente positivo (“ótimo”) em um contexto claramente negativo parece ter confundido o modelo.

Estes exemplos indicam que, assim como nas outras emoções analisadas, o sarcasmo e a ironia representam um desafio substancial para a classificação correta da emoção. Apesar disso, o desempenho geral do modelo nessa categoria pode ser considerado satisfatório, dada a complexidade da tarefa.

Tabela 6 – Exemplos de erros e acertos para Admiração

| | Acertos (VP) | Erros (FP) |
|---|--|--|
| 1 | Impressionante sua fala na reunião, presidente. Reeleito até 2038. Parabéns. Impressionante mesmo. Preocupação voltada aos cidadãos Brasileiro e ao Brasil. Grande dia. #fechadocombolsonaro | é corajoso para falar de fraude, para ofender repórter, para falar que não vai ter eleições. Mais peida para enfrentar as falas do #Respondebolsonaro os irmãos Miranda falaram ou não a verdade! #24JForaBolsonaro #LulaCuidaDoPovo #BolsonaroCorrupto |
| 2 | vc é o melhor presidente que essa porca república já viu | Parabéns por fazer o mínimo |
| 3 | Esse e nosso Presidente. Show de bola BOLSONARO . Está deixa só a mídia e a esquerdalha maluca e estamos gostando . | No mínimo preso ótimo . Que apanhe bastante antes e lá na cadeia de bônus . |

6 Conclusões

Este trabalho focou-se na utilização do GoEmotions para Português para análise das emoções presentes nos *tweets* de resposta aos posts do ex-presidente do Brasil, Jair Messias Bolsonaro. Para isso, foi utilizado o dataset fornecido pelo *GATE Team*, e extraídos os *tweets* de interesse para a pesquisa. Este processo resultou em um corpus final de aproximadamente 5 milhões de *tweets*, que passou por uma etapa de pré-processamento prévio à sua submissão ao GoEmotions. Este modelo foi responsável por associar cada *tweet* a uma emoção, entre as 27 opções disponíveis, ou categorizá-los como neutros.

A análise exploratória conduzida nesta pesquisa revelou um desequilíbrio substancial em relação à presença das diferentes emoções no corpus, em que a emoção mais predominante, “raiva”, se faz presente aproximadamente 247 vezes mais do que a emoção mais rara, “percepção”. As emoções “raiva”, “curiosidade” e “admiração”, que são as mais proeminentes no corpus, foram posteriormente submetidas a uma análise de erros. Este procedimento teve como objetivo avaliar a precisão e a qualidade das inferências realizadas pelo modelo GoEmotions, assim como identificar possíveis motivos para os erros e acertos do modelo.

No que se refere à identificação de emoções, o modelo apresentou performance variada. A emoção “raiva” foi identificada de forma eficaz, provavelmente atribuída à presença de palavras inflamatórias, manifestando uma precisão consideravelmente superior ao valor esperado inicialmente. O desempenho para a emoção “admiração”, embora razoável, mostrou que o modelo tem dificuldades em lidar com a complexidade linguística, em particular quando envolve ironia ou sarcasmo. A emoção “curiosidade”, por outro lado, revelou desafios significativos para o modelo, sobretudo devido à dificuldade em discernir perguntas genuínas de perguntas retóricas ou carregadas de ironia e sarcasmo.

A presente pesquisa demonstrou o potencial da aplicação de modelos de PLN na análise de emoções em *tweets* no contexto da política brasileira, ao mesmo tempo que revelou limitações que merecem ser objeto de futuras pesquisas. Os resultados enfatizam a necessidade de aprimoramentos contínuos na capacidade desses modelos para lidar com nuances de linguagem, como sarcasmo e ironia.

6.1 Trabalhos Futuros

Em vista dos resultados obtidos e das observações realizadas durante o desenvolvimento desta pesquisa, foram identificadas diversas oportunidades para estudos subsequentes. As seguintes sugestões são apresentadas como possíveis direções para futuros trabalhos:

- **Aprimoramento da capacidade de modelos de PLN em lidar com ironia e sarcasmo:** Dado que esta foi uma dificuldade identificada durante a análise, futuros trabalhos podem se concentrar em aprimorar essa habilidade, treinando modelos capazes de entender melhor as nuances da língua portuguesa;
- **Ampliação do espectro de emoções analisadas:** Dado o desequilíbrio observado nas instâncias de emoções no corpus, seria interessante investigar outras emoções menos prevalentes e entender o comportamento do modelo nessas categorias;
- **Determinação de um limiar seguro para a classificação de cada emoção:** Outro caminho promissor para pesquisas futuras pode envolver a tentativa de determinar um limiar seguro para a classificação de cada emoção. Por exemplo, considerar a aplicação de um score maior ou igual à média poderia potencialmente aumentar a precisão da classificação das emoções.
- **Desenvolvimento de uma taxonomia de emoções com base na cultura brasileira e treinamento de modelo de AS:** visto que uma taxonomia de emoções depende muito do contexto cultural e linguístico em que foi desenvolvida, uma sugestão é a elaboração de uma taxonomia de emoções específica para o contexto brasileiro, levando em conta as nuances culturais, linguísticas e emocionais que são exclusivas ao país. Uma vez desenvolvida essa taxonomia, ela poderia servir como base para treinar um modelo de AS adaptado às particularidades do Brasil.
- **Integração com as Ciências Sociais e Políticas:** Estudos futuros poderiam incorporar uma abordagem interdisciplinar, combinando análise de sentimentos com teorias das ciências sociais e políticas para proporcionar uma compreensão mais aprofundada do impacto das emoções na formação de opiniões e comportamentos políticos. Isso poderia envolver a análise de como as emoções expressas nos *tweets* estão correlacionadas com eventos políticos específicos, mudanças no apoio aos partidos ou políticos, entre outros;
- **Análise temporal das emoções:** investigar a evolução das emoções ao longo do tempo. Isso poderia permitir a identificação de tendências ou padrões nas emoções expressas em resposta a eventos ou decisões políticas específicas. Tal análise poderia se beneficiar do cruzamento de dados com outras áreas de estudo, como a ciência política, para melhor entender como os sentimentos da população evoluem ao longo do tempo em resposta a diferentes situações e contextos;
- **Expansão do contexto de análise:** O modelo GoEmotions para português foi testado neste estudo em um contexto específico, o da política brasileira. No entanto, a inferência e análise das emoções poderia ser expandida para outras áreas, como *review* de produtos ou em posts de redes sociais não relacionados à política.

Referências

- AJZENMAN, N.; CAVALCANTI, T.; MATA, D. D. More than words: Leaders' speech and risky behavior during a pandemic. *Cambridge Working Papers in Economics*, Faculty of Economics, University of Cambridge, 2020. Citado na página 21.
- ALPAYDIN, E. *Introduction to Machine Learning*. 2nd. ed. [S.l.]: The MIT Press, 2010. ISBN 026201243X. Citado na página 24.
- BAHDANAU, D.; CHO, K.; BENGIO, Y. *Neural Machine Translation by Jointly Learning to Align and Translate*. arXiv, 2014. Disponível em: <<https://arxiv.org/abs/1409.0473>>. Citado na página 31.
- BAHDANAU, D.; CHO, K.; BENGIO, Y. *Neural Machine Translation by Jointly Learning to Align and Translate*. 2016. Citado na página 30.
- BENGIO, Y. et al. A neural probabilistic language model. *J. Mach. Learn. Res.*, JMLR.org, v. 3, n. null, p. 1137–1155, mar 2003. ISSN 1532-4435. Citado na página 30.
- BHARATHI, S.; GEETHA, A. Sentiment analysis for effective stock market prediction. *International Journal of Intelligent Engineering and Systems*, v. 10, p. 146–154, 2017. Citado na página 19.
- BLEI, D. M.; NG, A. Y.; JORDAN, M. I. Latent dirichlet allocation. *J. Mach. Learn. Res.*, JMLR.org, v. 3, n. null, p. 993–1022, mar 2003. ISSN 1532-4435. Citado na página 45.
- BOJANOWSKI, P. et al. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, MIT Press, Cambridge, MA, v. 5, p. 135–146, 2017. Disponível em: <<https://aclanthology.org/Q17-1010>>. Citado na página 30.
- BROWN, T. B. et al. *Language Models are Few-Shot Learners*. arXiv, 2020. Disponível em: <<https://arxiv.org/abs/2005.14165>>. Citado na página 32.
- CARVALHO, P.; SILVA, M. J. Sentilex-pt: Principais características e potencialidades. *Oslo Studies in Language*, v. 7, 2015. Disponível em: <<https://api.semanticscholar.org/CorpusID:61390164>>. Citado na página 45.
- CHAWLA, N. V. et al. Smote: Synthetic minority over-sampling technique. *J. Artif. Int. Res.*, AI Access Foundation, El Segundo, CA, USA, v. 16, n. 1, p. 321–357, jun 2002. ISSN 1076-9757. Citado na página 27.
- CHURCH, K. W. Word2vec. *Natural Language Engineering*, Cambridge University Press, v. 23, n. 1, p. 155–162, 2017. Citado na página 44.
- COHEN, A. et al. *Maximum Entropy Diverse Exploration: Disentangling Maximum Entropy Reinforcement Learning*. 2019. Citado na página 46.

- COPPERSMITH, G.; DREDZE, M.; HARMAN, C. Quantifying mental health signals in Twitter. In: *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*. Baltimore, Maryland, USA: Association for Computational Linguistics, 2014. p. 51–60. Disponível em: <<https://www.aclweb.org/anthology/W14-3207>>. Citado na página 20.
- CORTES, C.; VAPNIK, V. Support-vector networks. *Machine learning*, Springer, v. 20, n. 3, p. 273–297, 1995. Citado 2 vezes nas páginas 46 e 47.
- COWEN, A. S.; KELTNER, D. Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proceedings of the National Academy of Sciences*, v. 114, n. 38, p. E7900–E7909, 2017. Disponível em: <<https://www.pnas.org/doi/abs/10.1073/pnas.1702247114>>. Citado na página 42.
- CRISTIANI, A.; LIEIRA, D.; CAMARGO, H. A sentiment analysis of brazilian elections tweets. In: *Anais do VIII Symposium on Knowledge Discovery, Mining and Learning*. Porto Alegre, RS, Brasil: SBC, 2020. p. 153–160. ISSN 2763-8944. Disponível em: <<https://sol.sbc.org.br/index.php/kdmile/article/view/11971>>. Citado 2 vezes nas páginas 47 e 48.
- CUI, Y. et al. Class-balanced loss based on effective number of samples. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2019. p. 9268–9277. Citado 2 vezes nas páginas 28 e 55.
- DEMSZKY, D. et al. GoEmotions: A dataset of fine-grained emotions. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics, 2020. p. 4040–4054. Disponível em: <<https://aclanthology.org/2020.acl-main.372>>. Citado 3 vezes nas páginas 19, 42 e 44.
- DEVLIN, J. et al. BERT: Pre-training of deep bidirectional transformers for language understanding. In: *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://www.aclweb.org/anthology/N19-1423>>. Citado na página 32.
- DOMINGOS, P. A few useful things to know about machine learning. *Commun. ACM*, Association for Computing Machinery, New York, NY, USA, v. 55, n. 10, p. 78–87, oct 2012. ISSN 0001-0782. Disponível em: <<https://doi.org/10.1145/2347736.2347755>>. Citado na página 25.
- EKMAN, P. Are there basic emotions? *Psychological review*, v. 99 3, p. 550–3, 1992. Citado 2 vezes nas páginas 41 e 42.
- FONTAINE, J. et al. Cognitive structure of emotion terms in indonesia and the netherlands. *Cognition Emotion*, January 1, p. 61–86, 09 2010. Citado na página 39.
- F.R.S., K. P. Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, Taylor Francis, v. 2, n. 11, p. 559–572, 1901. Citado na página 25.

GAYATHRI, P.; GOPICHAND, G.; AMALI, S. H. G. B. A case study on artificial intelligence application in medical diagnostics. *WWJMRD*, v. 4, n. 12, p. 100–103, 2018. Citado na página 27.

GOMEZ-URIBE, C. A.; HUNT, N. The netflix recommender system: Algorithms, business value, and innovation. *ACM Trans. Manage. Inf. Syst.*, Association for Computing Machinery, New York, NY, USA, v. 6, n. 4, dec 2016. ISSN 2158-656X. Disponível em: <<https://doi.org/10.1145/2843948>>. Citado na página 19.

GORMLEY, C.; TONG, Z. *Elasticsearch: The Definitive Guide*. 1st. ed. [S.l.]: O'Reilly Media, Inc., 2015. ISBN 1449358543. Citado na página 51.

GUDIKANDULA, P. *A Beginner Intro to Neural Networks*. 2019. Disponível em: <<https://purnasaigudikandula.medium.com/a-beginner-intro-to-neural-networks-543267bda3c8>>. Citado na página 29.

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>>. Citado 5 vezes nas páginas 20, 28, 54, 55 e 56.

HAYKIN, S. *Neural Networks and Learning Machines*. Pearson, 2009. (Pearson International Edition). ISBN 9780131293762. Disponível em: <<https://books.google.com.br/books?id=KCwWOAAACAAJ>>. Citado na página 28.

HE, H.; GARCIA, E. A. Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, v. 21, n. 9, p. 1263–1284, 2009. Citado na página 27.

HO, T. K. Random decision forests. In: IEEE. *Proceedings of 3rd international conference on document analysis and recognition*. [S.l.], 1995. v. 1, p. 278–282. Citado na página 24.

IBANEZ, D. *Encoder-Decoder Models for Natural Language Processing*. 2022. Disponível em: <<https://www.baeldung.com/cs/nlp-encoder-decoder-models>>. Citado 2 vezes nas páginas 29 e 33.

IYYER, M. et al. Political ideology detection using recursive neural networks. In: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Baltimore, Maryland: Association for Computational Linguistics, 2014. p. 1113–1122. Disponível em: <<https://www.aclweb.org/anthology/P14-1105>>. Citado na página 20.

JAPKOWICZ, N.; SHAH, M. *Evaluating Learning Algorithms: A Classification Perspective*. USA: Cambridge University Press, 2011. ISBN 0521196000. Citado na página 34.

KIBRIYA, A. M. et al. Multinomial naive bayes for text categorization revisited. In: WEBB, G. I.; YU, X. (Ed.). *AI 2004: Advances in Artificial Intelligence*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005. p. 488–499. ISBN 978-3-540-30549-1. Citado na página 46.

- KIBUUKA, B. Complicity and synergy between bolsonaro and brazilian evangelicals in covid-19 times: Adherence to scientific negationism for political-religious reasons. *International Journal of Latin American Religions*, Faculty of Economics, University of Cambridge, v. 4, p. 288–317, December 2020. ISSN 2509-9665. Citado na página 21.
- KRISHNA, P. P.; SHARADA, A. Word embeddings - skip gram model. In: GUNJAN, V. K. et al. (Ed.). *ICICCT 2019 – System Reliability, Quality Control, Safety, Maintenance and Management*. Singapore: Springer Singapore, 2020. p. 133–139. ISBN 978-981-13-8461-5. Citado na página 44.
- LIU, X. et al. Emotion classification for short texts: an improved multi-label method. *Humanit Soc Sci Commun*, Nature Publishing Group, v. 10, p. 306, 2023. Disponível em: <<https://doi.org/10.1057/s41599-023-01816-6>>. Citado na página 30.
- LLOYD, S. Least squares quantization in pcm. *IEEE Transactions on Information Theory*, v. 28, n. 2, p. 129–137, 1982. Citado na página 25.
- LOPER, E.; BIRD, S. *NLTK: The Natural Language Toolkit*. arXiv, 2002. Disponível em: <<https://arxiv.org/abs/cs/0205028>>. Citado na página 54.
- LUGER, G. F.; STUBBLEFIELD, W. A. *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*. 3rd. ed. USA: Addison-Wesley Longman Publishing Co., Inc., 1997. ISBN 0805311963. Citado na página 28.
- MAYER, J. H. et al. Towards natural language processing: An accounting case study. In: *ICIS 2020 Proceedings*. [s.n.], 2020. Disponível em: <https://aisel.aisnet.org/icis2020/practice_is_research/practice_is_research/1>. Citado na página 27.
- MEDHAT, W.; HASSAN, A.; KORASHY, H. Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, v. 5, n. 4, p. 1093–1113, 2014. ISSN 2090-4479. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2090447914000550>>. Citado 2 vezes nas páginas 37 e 38.
- MNASRI, M. Recent advances in conversational NLP : Towards the standardization of chatbot building. *CoRR*, abs/1903.09025, 2019. Disponível em: <<http://arxiv.org/abs/1903.09025>>. Citado na página 27.
- MOFIJUR, M. et al. Impact of covid-19 on the social, economic, environmental and energy domains: Lessons learnt from a global pandemic. *Sustainable Production and Consumption*, v. 26, p. 343–359, 2021. ISSN 2352-5509. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2352550920313567>>. Citado na página 20.
- NADKARNI, P. M.; OHNO-MACHADO, L.; CHAPMAN, W. W. Natural language processing: an introduction. *Journal of the American Medical Informatics Association*, v. 18, n. 5, p. 544–551, 09 2011. ISSN 1067-5027. Disponível em: <<https://doi.org/10.1136/amiajnl-2011-000464>>. Citado na página 25.
- PANDYA, R. V. R. *Generalized Attention Mechanism and Relative Position for Transformer*. arXiv, 2022. Disponível em: <<https://arxiv.org/abs/2208.10247>>. Citado na página 32.

- PREOTIUC-PIETRO, D. et al. Studying the dark triad of personality through twitter behavior. In: *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. New York, NY, USA: Association for Computing Machinery, 2016. (CIKM '16), p. 761–770. ISBN 9781450340731. Disponível em: <<https://doi.org/10.1145/2983323.2983822>>. Citado na página 20.
- RADFORD, A. et al. Language models are unsupervised multitask learners. In: . [S.l.: s.n.], 2019. Citado 2 vezes nas páginas 19 e 30.
- REYES, A.; ROSSO, P.; VEALE, T. A multidimensional approach for detecting irony in twitter. *Language Resources and Evaluation*, v. 47, 03 2013. Citado na página 20.
- REZENDE, S. *Sistemas inteligentes: fundamentos e aplicações*. Manole, 2003. ISBN 9788520416839. Disponível em: <https://books.google.com.br/books?id=UsJe_PlbnWcC>. Citado na página 24.
- RUSSELL, J.; LEWICKA, M.; NIIT, T. A cross-cultural study of a circumplex model of affect. *Journal of Personality and Social Psychology*, v. 57, p. 848–856, 11 1989. Citado na página 39.
- RUSSELL, J. A. A circumplex model of affect. *Journal of Personality and Social Psychology*, v. 39, n. 6, p. 1161–1178, 1980. Citado na página 39.
- RUSSELL, J. A. Pancultural aspects of the human conceptual organization of emotions. *Journal of Personality and Social Psychology*, v. 45, p. 1281–1288, 1983. Disponível em: <<https://api.semanticscholar.org/CorpusID:17508583>>. Citado na página 39.
- RUSSELL, S.; NORVIG, P. *Artificial Intelligence: A Modern Approach*. 3rd. ed. USA: Prentice Hall Press, 2009. ISBN 0136042597. Citado na página 23.
- SALUNKE, T.; NICHITE, U. Recommender systems in e-commerce. Unpublished, 2022. Disponível em: <<https://rgdoi.net/10.13140/RG.2.2.10194.43202>>. Citado na página 19.
- SANTANA, B. S.; VANIN, A. A. Detecting group beliefs related to 2018's brazilian elections in tweets A combined study on modeling topics and sentiment analysis. *CoRR*, abs/2006.00490, 2020. Disponível em: <<https://arxiv.org/abs/2006.00490>>. Citado 2 vezes nas páginas 44 e 45.
- SCHERER, K. R. Emotion as a multicomponent process: A model and some cross-cultural data. *Review of Personality & Social Psychology*, v. 5, p. 37–63, 1984. Citado na página 39.
- SCHERER, K. R. What are emotions? and how can they be measured? *Social Science Information*, v. 44, n. 4, p. 695–729, 2005. Disponível em: <<https://doi.org/10.1177/0539018405058216>>. Citado 2 vezes nas páginas 39 e 40.
- SENNRICH, R.; HADDOW, B.; BIRCH, A. *Neural Machine Translation of Rare Words with Subword Units*. arXiv, 2015. Disponível em: <<https://arxiv.org/abs/1508.07909>>. Citado na página 31.
- SHAVER, P. et al. Emotion knowledge: further exploration of a prototype approach. *Journal of personality and social psychology*, v. 52, n. 6, p. 1061–1086, June 1987. ISSN 0022-3514. Disponível em: <<https://doi.org/10.1037//0022-3514.52.6.1061>>. Citado na página 39.

- SHAVER, P. R.; WU, S.; SCHWARTZ, J. C. Cross-cultural similarities and differences in emotion and its representation. In: . [s.n.], 1992. Disponível em: <<https://api.semanticscholar.org/CorpusID:140741300>>. Citado na página 39.
- SHU, K. et al. Fake news detection on social media: A data mining perspective. *SIGKDD Explor. Newsl.*, Association for Computing Machinery, New York, NY, USA, v. 19, n. 1, p. 22–36, sep 2017. ISSN 1931-0145. Disponível em: <<https://doi.org/10.1145/3137597.3137600>>. Citado na página 20.
- SOUZA, B. A. et al. For or against? polarity analysis in tweets about impeachment process of brazil president. In: *Proceedings of the 22nd Brazilian Symposium on Multimedia and the Web*. New York, NY, USA: Association for Computing Machinery, 2016. (Webmedia '16), p. 335–338. ISBN 9781450345125. Disponível em: <<https://doi.org/10.1145/2976796.2988216>>. Citado 2 vezes nas páginas 46 e 47.
- SOUZA, F.; NOGUEIRA, R.; LOTUFO, R. BERTimbau: pretrained BERT models for Brazilian Portuguese. In: *9th Brazilian Conference on Intelligent Systems, BRACIS, Rio Grande do Sul, Brazil, October 20-23 (to appear)*. [S.l.: s.n.], 2020. Citado na página 33.
- SUTSKEVER, I.; VINYALS, O.; LE, Q. V. *Sequence to Sequence Learning with Neural Networks*. 2014. Citado na página 29.
- TRNKA, M. et al. Mapping discrete emotions in the dimensional space: An acoustic approach. *Electronics*, v. 10, n. 23, 2021. ISSN 2079-9292. Disponível em: <<https://www.mdpi.com/2079-9292/10/23/2950>>. Citado na página 39.
- TUMASJAN, A. et al. Predicting elections with twitter: What 140 characters reveal about political sentiment. *Proceedings of the International AAAI Conference on Web and Social Media*, v. 4, n. 1, p. 178–185, May 2010. Disponível em: <<https://ojs.aaai.org/index.php/ICWSM/article/view/14009>>. Citado na página 20.
- VASWANI, A. et al. *Attention Is All You Need*. arXiv, 2017. Disponível em: <<https://arxiv.org/abs/1706.03762>>. Citado na página 32.
- WEBB, G. I. Naïve bayes. In: _____. *Encyclopedia of Machine Learning*. Boston, MA: Springer US, 2010. p. 713–714. ISBN 978-0-387-30164-8. Disponível em: <https://doi.org/10.1007/978-0-387-30164-8_576>. Citado na página 47.
- WU, X. et al. Top 10 algorithms in data mining. *Knowledge and information systems*, Springer, v. 14, n. 1, p. 1–37, 2008. Citado na página 24.
- WU, Y. et al. *Google’s Neural Machine Translation System: Bridging the Gap between Human and Machine Translation*. arXiv, 2016. Disponível em: <<https://arxiv.org/abs/1609.08144>>. Citado na página 31.
- ZHANG, L.; LIU, B. Sentiment analysis and opinion mining. In: *Encyclopedia of Machine Learning and Data Mining*. [s.n.], 2012. Disponível em: <<https://api.semanticscholar.org/CorpusID:38022159>>. Citado na página 30.
- ZHU, X.; GOLDBERG, A. B. Introduction to semi-supervised learning. In: *Introduction to Semi-Supervised Learning*. [S.l.: s.n.], 2009. Citado na página 25.