

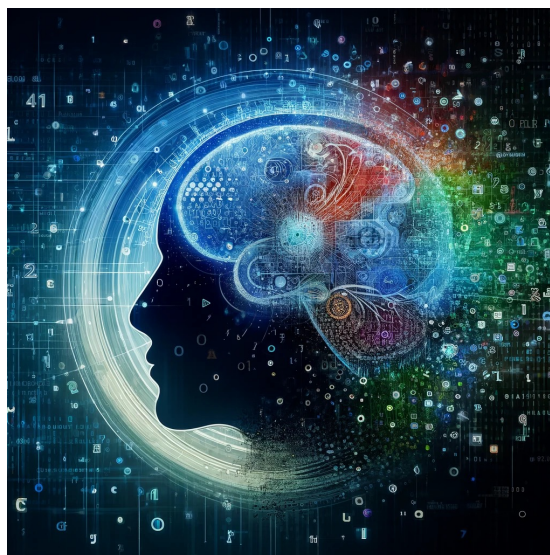
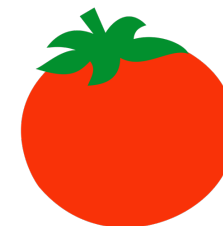
**MBA  
USP  
ESALQ**

# **Comparação entre modelos para Análise de Sentimentos no Contexto de Avaliações de Filmes**

Discente: Rodrigo Franciozi Rodrigues da Silva  
Orientadora: Adâmara Santos Gonçalves Felício



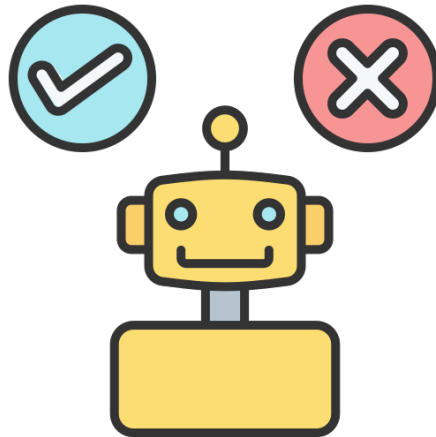
A indústria de áudio visual, em constante crescimento, busca entender as percepções do público para criar obras de sucesso. A análise de comentários e avaliações de plataformas como IMDB, Rotten Tomatoes e Metacritic, além das redes sociais, tornou-se fundamental para medir o engajamento e potencial sucesso das produções.



A análise de sentimentos, ou mineração de opiniões, utiliza técnicas de inteligência artificial para avaliar polaridades de sentimentos em textos. Esta abordagem permite transformar dados não estruturados em informações úteis para a tomada de decisões, destacando-se nas áreas de PLN [Processamento de Linguagem Natural].



O objetivo do trabalho foi investigar como modelos de análise de sentimentos podem efetivamente distinguir opiniões **positivas**, **negativas** ou **nêutras**, com o principal intuito de se comparar a eficácia de modelos convencionais frente aos de aprendizado profundo, baseados em redes neurais.



ML Clássico

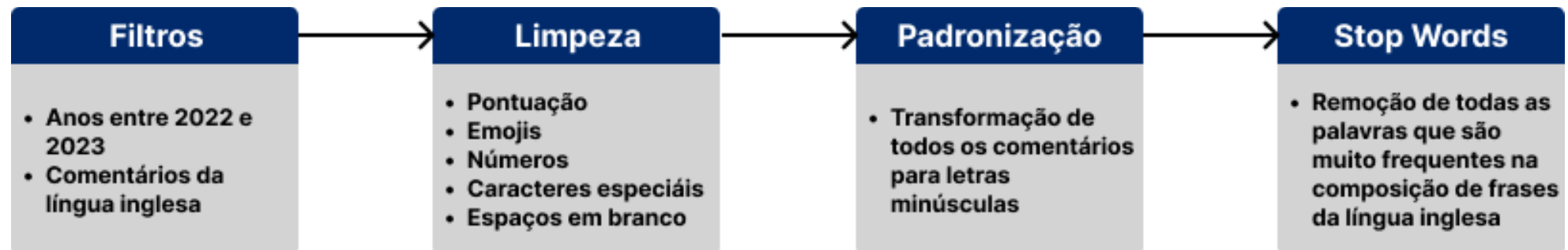
Comparação



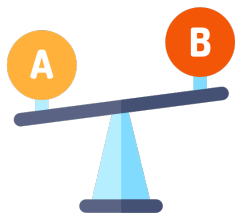
Rede Neural

## Webscrapping

- Técnica de raspagem de dados feita diretamente no site do IMDB (Mine Çetinkaya-Rundel e Mine Dogucu, 2021).
- Bastante útil quando não se tem acesso aos dados por meio de API;



Fonte: Dados originais da pesquisa



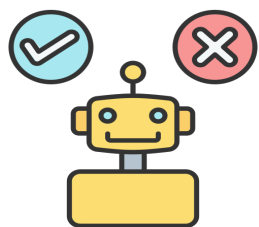
Aplicação de técnica de balanceamento de classes (Nassera *et al.*, 2023).



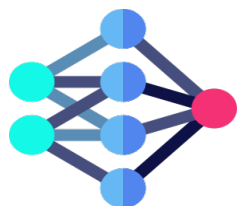
Divisão de 80/20 para treino e teste seguindo o conceito da regra de Pareto (Rosie *et al.*, 2014).



\*Modelo pré-treinado pela Huggin face



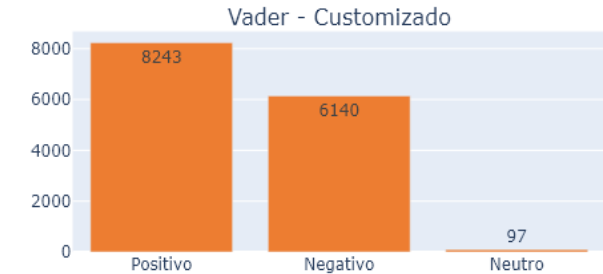
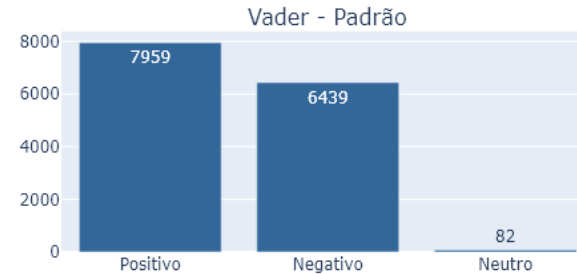
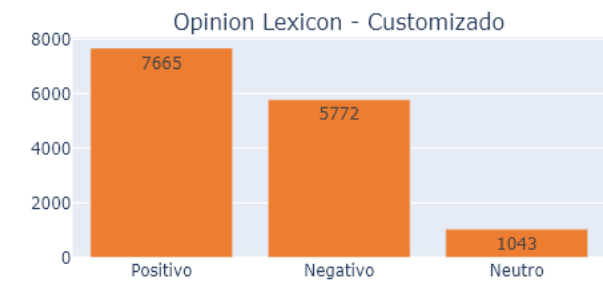
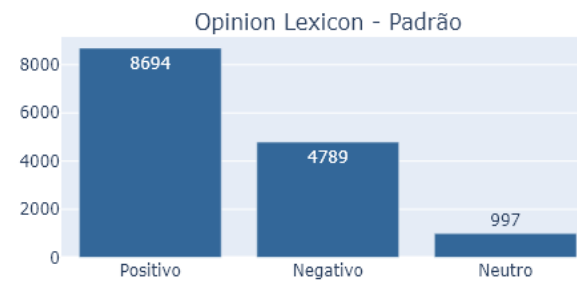
Treinamento do modelo Naïve Bayes multinomial (3 classes), que calcula a probabilidade de um documento (frase) pertencer a determinada classe, conforme a frequência da contagem de palavras no documento (Christine *et al.*, 2023).



Re-treinamento<sup>\*</sup> do modelo de rede neural BERT [Bidirectional Encoder Representations from Transformers] que é considerado o estado da arte para classificação devido a sua capacidade de considerar o contexto das palavras em uma frase e como seu significado pode mudar (Jacob *et al.* 2018) .

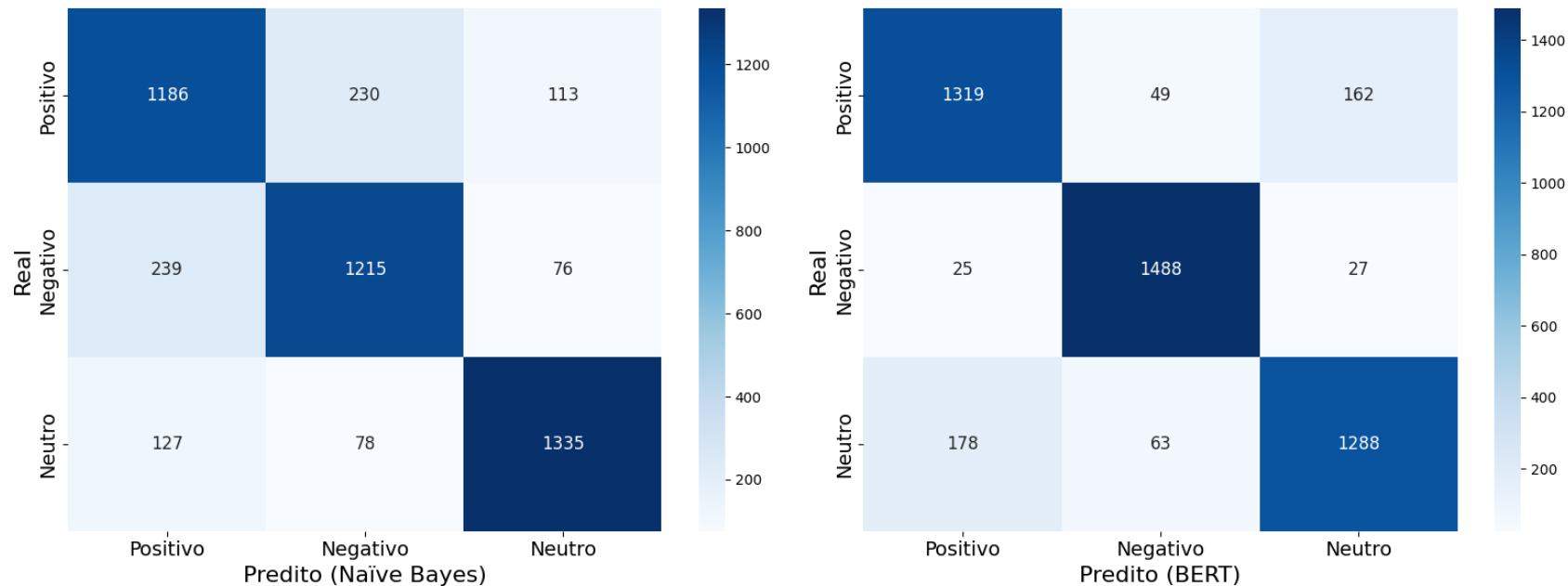
## Mas e as labels para a aplicação do aprendizado supervisionado?

1. Aplicação de uma técnica **não supervisionada** que atribui um valor (peso) de acordo com um dicionário de palavras previamente definido (Fabrício *et al.*, 2015);
2. VADER x Opinion Lexicon;
3. Opinion Lexicon foi o escolhido por apresentar uma maior sensibilidade na classificação das frases neutras e por ter sua origem voltada para avaliações de sites de compra;



Fonte: Resultados originais da pesquisa

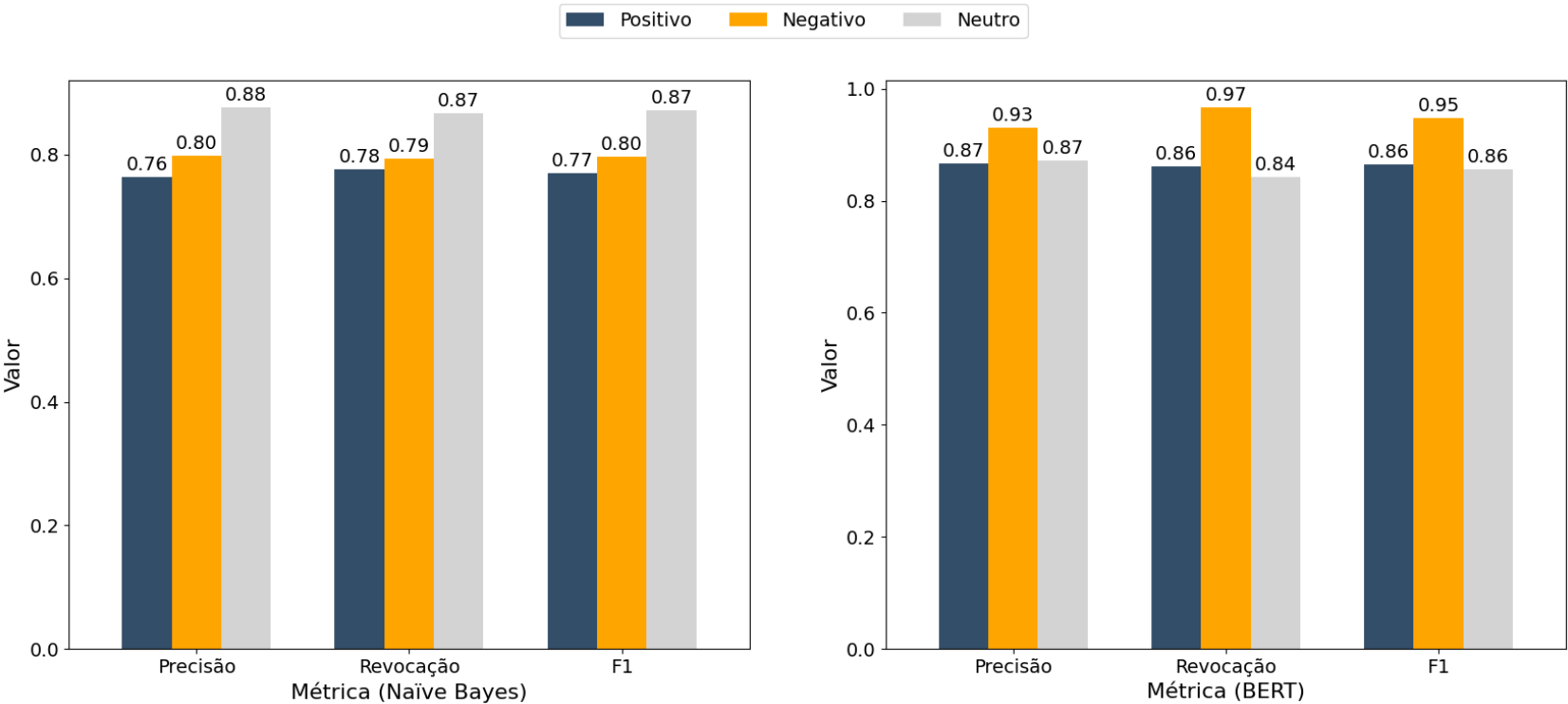
- As matrizes revelam um desempenho melhor no modelo BERT para classificações de sentimentos positivos e negativos;
- A acurácia para os modelos Naïve Bayes e BERT foi de 81% e 89% respectivamente. Uma performance ligeiramente maior, mas consistente com a natureza do algoritmo de aprendizado profundo;
- A escolha do modelo BERT mostra uma particularidade vantajosa em cenários onde a distinção precisa entre comentários positivos e negativos, teria uma maior importância.



Fonte: Resultados originais da pesquisa

# Resultados – Métricas de classificação

- **Precisão:** Reflete a porcentagem de identificações corretas de uma classe, considerando as previsões;
- **Revocação:** Mede a capacidade do modelo de encontrar as instâncias relevantes de uma determinada classe;
- **F1:** Média harmônica entre a Precisão e a Revocação, fornecendo valores que equilibram o fato de identificar corretamente um classe e não perder muitas instâncias daquela classe;
- **BERT** se sobressai em praticamente todas as métricas, sugerindo uma maior aptidão para entender o contexto dos comentários e classificá-los corretamente.



Fonte: Resultados originais da pesquisa



- ✓ Apesar da simplicidade do modelo Naïve Bayes, o mesmo obteve um desempenho ligeiramente maior na predição das classificações Neutras.
- ✓ O modelo BERT superou significativamente em precisão e revocação, beneficiando-se da capacidade de melhor compreensão das nuances da linguagem.
- ✓ Reforça o quanto é relevante a implementação do aprendizado profundo em tarefas de processamento de linguagem natural, especialmente no âmbito de análise de sentimentos.
- ✓ A escolha dos modelos deve considerar o contexto da aplicação. Enquanto o BERT é adequado para cenários onde exigem altas precisões e contextualização, o Naïve Bayes pode ser preferível para aplicações com limitações de recursos computacionais.
- ✓ As percepções desse estudo são relevantes para empresas que desejam aplicar análises de sentimentos para poder direcionar estratégias de marketing baseadas em dados.

## Limitações

1. Recursos computacionais para treinamento do modelo BERT, o que impôs restrições ao tamanho do conjunto de dados e a relação da complexidade dos experimentos;
2. A natureza dos comentários ser apenas da língua inglesa, o que não necessariamente captou nuances culturais e linguísticas presentes em avaliações de outras regiões;
3. Restrição monetária de acesso à API do IMDB;

## Melhorias

1. Proposta de exploração de modelos híbridos, visando otimizar tanto a precisão quanto a eficiência computacional;
2. Explorar o desempenho de outros modelos de aprendizado de máquina;
3. A investigação do desempenho dos modelos levando em consideração as bases desbalanceadas, bem como a aplicação de técnicas mais avançadas;

Christine Dewi, Rung-Ching Chen, Henochi Juli Cristiano, Francesco Cauteruccio. Multinomial Naïve Bayes Classifier for Sentiment Analysis of Internet Movie Database. Vietnam Journal of Computer Science, 2023.

Fabício Bevenuto, Filipe Ribeiro, Matheus Araujo. Métodos para Análises de Sentimentos em Mídias Sociais. Curso rápido em uma conferência de webmedia, 2015.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Google AI Language, 2018.

Mine Çetinkaya-Rundel, Mine Dogucu. Web Scraping in the Statistics and Data Science Curriculum: Challenges and Opportunities, Journal of Statistics and Data Science Education, 2021.

Nassera Habbat, Hicham Nouri, Houda Anoun, Larbi Hassouni. Sentiment analysis of imbalanced datasets using BERT and ensemble stacking for deep learning. Research Laboratory on New Economy and Development, 2023.

Rosie Dunfornd, Quanrong Su, Ekraj Tamang, Abigail Wintour. The Pareto Principle. The Plymouth Student Scientist, 2024.

Muito obrigado pela atenção!