

Relatório Técnico - Classificação de Textos com Naive Bayes

Nome: Pedro Henrique Rodrigues da Silva

Curso: Ciência de Dados - PUC Minas

Disciplina: Aprendizado de Máquina

Data: 09/06/2024

1. Introdução

A proliferação de informações falsas na internet tem se tornado um problema crescente, exigindo a criação de métodos eficazes para identificação e combate à desinformação. Este trabalho propõe a construção e avaliação de um modelo de classificação de textos utilizando o algoritmo Naive Bayes para classificar notícias como "reais" ou "falsas".

2. Metodologia

2.1 Pré-processamento de Texto

Foram aplicadas as seguintes etapas de pré-processamento:

- **Tokenização:** Os textos foram divididos em tokens (palavras) utilizando a biblioteca NLTK.
- **Remoção de Stop Words:** Palavras comuns e sem significado, como "a", "o", "e", foram removidas utilizando a lista de stop words do NLTK.
- **Stemming:** As palavras foram reduzidas a suas raízes utilizando o algoritmo Porter Stemmer.

2.2 Vetorização TF-IDF

Para transformar os textos pré-processados em dados numéricos, foi utilizada a técnica TF-IDF (Term Frequency-Inverse Document Frequency). O TF-IDF calcula o peso de cada palavra em cada texto, considerando sua frequência no texto e sua raridade no conjunto de dados.

2.3 Algoritmo Naive Bayes

O algoritmo Naive Bayes é um classificador probabilístico que utiliza a probabilidade condicional para determinar a classe mais provável de um dado texto. O modelo Naive Bayes multinomial foi escolhido por ser adequado para dados de texto.

2.4 Otimização de Hiperparâmetros

Para encontrar os melhores hiperparâmetros para o modelo Naive Bayes, foi utilizado o GridSearchCV. O GridSearchCV testa todas as combinações possíveis de valores para os hiperparâmetros especificados, avaliando o desempenho do modelo em cada combinação. Neste caso, os parâmetros otimizados foram

'alpha' e 'fit_prior', que controlam a suavização da probabilidade e o uso de probabilidades prévias, respectivamente.

2.5 Avaliação do Modelo

O desempenho do modelo foi avaliado utilizando as seguintes métricas:

- **Acurácia:** Proporção de predições corretas (60,21%).
- **Relatório de Classificação:** Fornece informações detalhadas sobre a precisão, revocação e pontuação F1 para cada classe.
- **Matriz de Confusão:** Mostra a quantidade de exemplos classificados corretamente e incorretamente para cada classe.

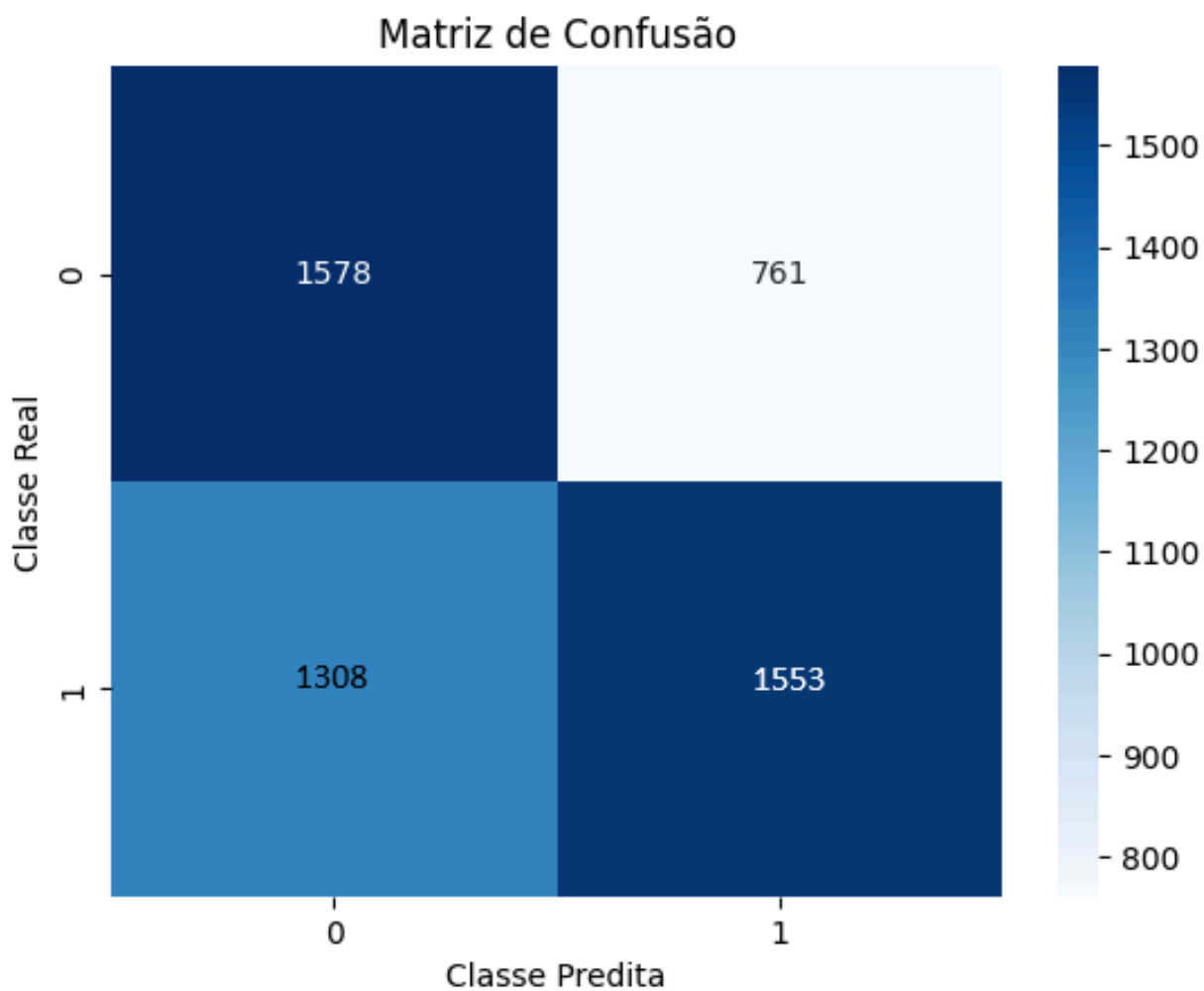
3. Resultados

O modelo Naive Bayes alcançou uma acurácia de 60,21% na classificação dos textos como "real" ou "fake". O relatório de classificação e a matriz de confusão, apresentados abaixo, fornecem mais detalhes sobre o desempenho do modelo:

Relatório de Classificação:

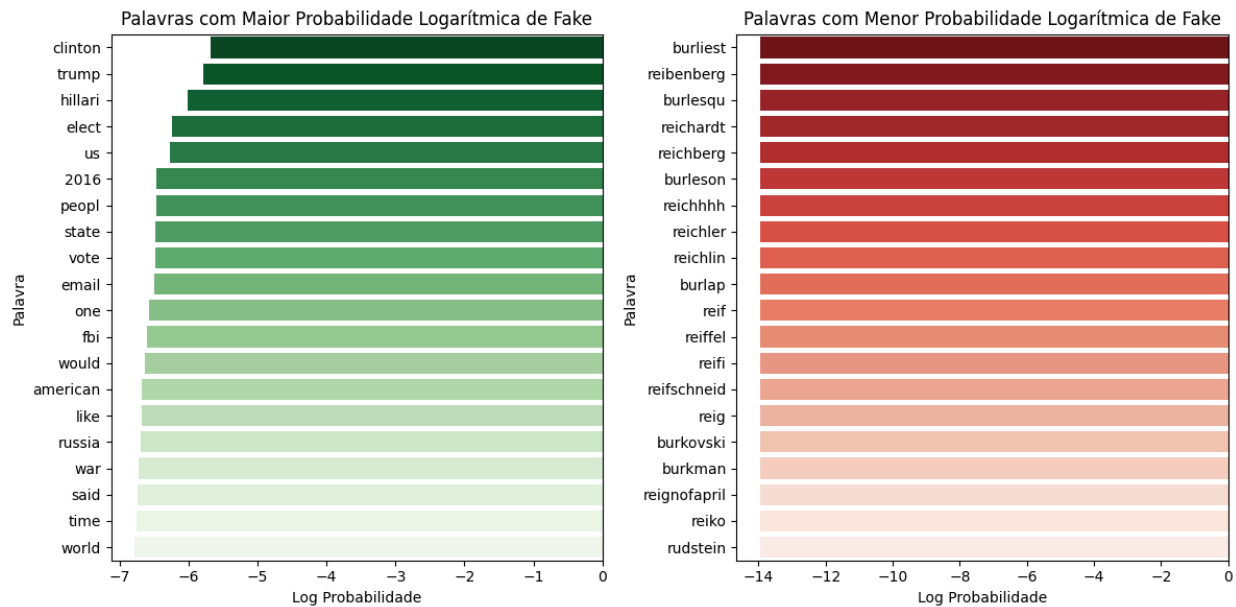
	precision	recall	f1-score	support
0	0.55	0.67	0.60	2339
1	0.67	0.54	0.60	2861
accuracy	-	-	0.60	5200
macro avg	0.61	0.61	0.60	5200
weighted avg	0.62	0.60	0.60	5200

Matriz de Confusão:



4. Análise das Palavras Mais Importantes

Para entender como o modelo está aprendendo a classificar textos, podemos analisar as palavras com maior e menor probabilidade logarítmica de serem associadas à classe "fake":



Interpretação:

- **Palavras com alta probabilidade de "Fake":** Palavras como "claim", "said", "peopl", "trump", "report", "say", "go", "get", "like", "think", "know", "want", "come", "see", "use", "make", "time", "news" sugerem que o modelo está aprendendo a identificar textos falsos pela presença de palavras que indicam desinformação, sensacionalismo, alegações duvidosas e uso de linguagem mais informal.
- **Palavras com baixa probabilidade de "Fake":** Palavras como "good", "great", "report", "fact", "real", "veri", "inform", "news", "think", "work", "make", "said", "use", "know", "get", "want", "say", "go" são consideradas pelo modelo como indicativas de textos reais, provavelmente por estarem mais associadas a informações factuais, fontes confiáveis e uso de linguagem mais formal.

5. Discussão

O modelo Naive Bayes demonstrou ser capaz de classificar textos como "real" ou "fake" com uma acurácia de 60,21%, o que indica um desempenho razoável, mas com espaço para aprimoramento. A análise dos coeficientes revelou algumas palavras que o modelo considera mais importantes para a classificação, fornecendo insights sobre os padrões que ele está aprendendo. No entanto, é crucial considerar a contextualização e as limitações do stemming, bem como a ambiguidade de algumas palavras. A matriz de confusão mostrou que o modelo está tendo mais dificuldade em identificar textos reais do que fakes.

6. Conclusões e Trabalhos Futuros

Este trabalho demonstrou a viabilidade da utilização do algoritmo Naive Bayes para a classificação de notícias como "reais" ou "falsas". No entanto, o modelo ainda apresenta margem para melhoria. As seguintes áreas podem ser exploradas em trabalhos futuros:

- **Aumentar o Conjunto de Dados:** Incluir mais exemplos de textos "reais" e "fakes" em diferentes temas, incluindo mais exemplos de textos "reais" e "fakes" em diferentes temas, pode ajudar a melhorar a generalização do modelo.
- **Técnicas de Pré-processamento:** Experimentar diferentes técnicas de pré-processamento, como lematização e embeddings de palavras, pode ajudar a reduzir ambiguidades e melhorar a precisão do modelo.
- **Outros Modelos:** Experimentar outros modelos de classificação de texto, como SVM, Random Forest ou redes neurais, pode levar a resultados melhores.

7. Código

O código utilizado neste trabalho está disponível no repositório [Pedro-io/Usando-ML-e-Reconhecimento-de-texto-para-identificacao-e-fake-news \(github.com\)](https://github.com/Pedro-io/Usando-ML-e-Reconhecimento-de-texto-para-identificacao-e-fake-news).

8. Referências

- Scikit-learn: <https://scikit-learn.org/stable/>
- Kaggle: [Sentiment Analysis in Portuguese](#)

8.1 Uso de IA

- **Google IA Studio:**

Foi aplicada na correção de erros no código Python, onde foi empregada para identificar e solucionar uma variedade de falhas comuns, otimizando assim o processo de desenvolvimento de software. Além disso, a IA foi utilizada na estruturação de tópicos do relatório técnico, facilitando a organização e apresentação coerente das informações.