

Processamento de linguagem natural

Análise de Sentimentos

Análise de sentimentos em reviews de produtos Amazon

Pedro Gabriel Santos do Couto Soares | pgscs@cin.ufpe.br
Carlos Eduardo Figueiredo Costa | cefc@cin.ufpe.br

Especialização
Deep Learning



Centro de
Informática
UFPE



UNIVERSIDADE
FEDERAL
DE PERNAMBUCO

Introdução

O contexto

1. **Cálculo correto de pontuações** para produtos após a venda:
 - Garante maior satisfação do cliente.
 - Destaca produtos de qualidade.
 - Melhora a experiência de compra.
2. **Ordenação correta dos comentários:**
 - Evita destaque de comentários enganosos.
 - Previne perdas financeiras e insatisfação do cliente.

Impacto esperado:

- **E-commerce:** Aumento das vendas e da confiança dos clientes.
- **Sellers:** Maior visibilidade para produtos bem avaliados.
- **Clientes:** Uma jornada de compra mais confiável e satisfatória.

Objetivos e metodologia

Objetivo:

- Classificar reviews em sentimentos (positivo, negativo e neutro).
- Comparar diferentes abordagens de modelagem.

Metodologia:

- Pré-processamento de dados.
- Análise exploratória e visualizações.
- Modelagem usando SVM (com Bag of Words e Embeddings) e BERT.

Dataset

- Kaggle dataset: amazon reviews for sentiment analysis
- url:
<https://www.kaggle.com/datasets/tarkkaanko/amazon>
- **4914** reviews

kaggle

Pré-processamento

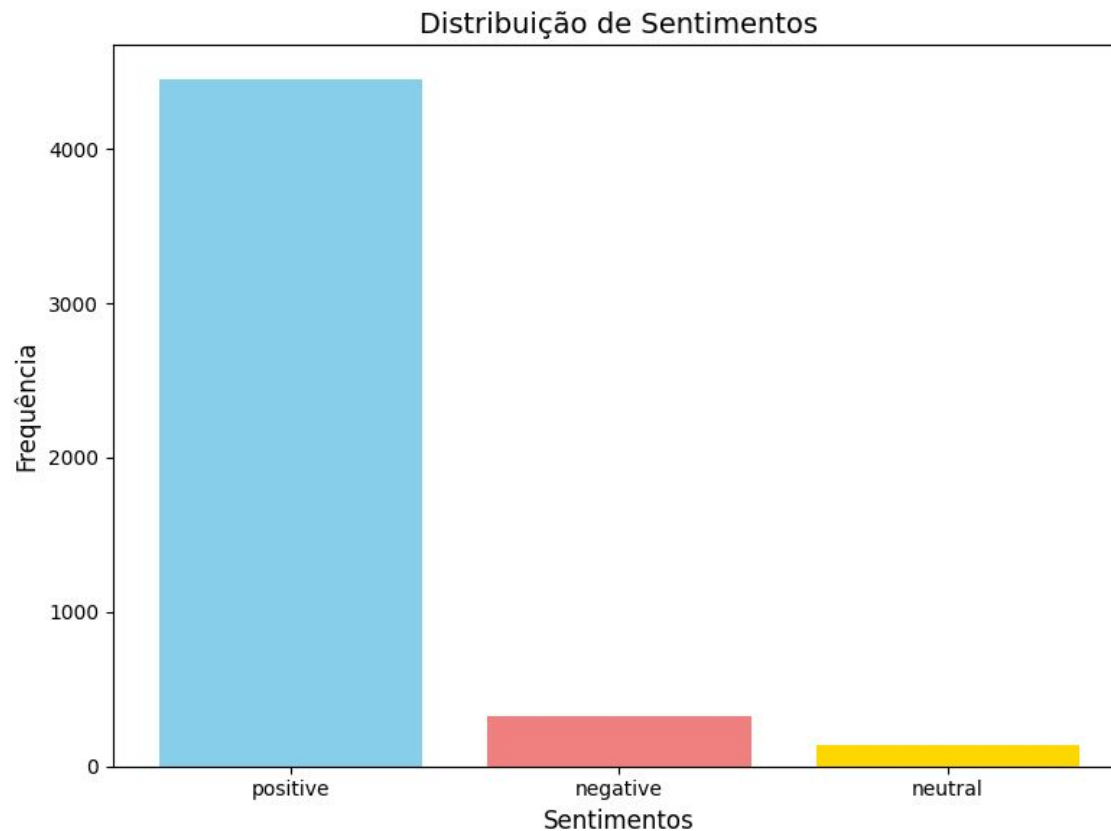
- Limpeza dos textos:
 - Remoção de pontuações e stopwords.
 - Conversão para minúsculas.
 - Remoção de reviews sem texto
- Mapeamento das notas em sentimentos.

Análise Exploratória - Nuvem de palavras geral

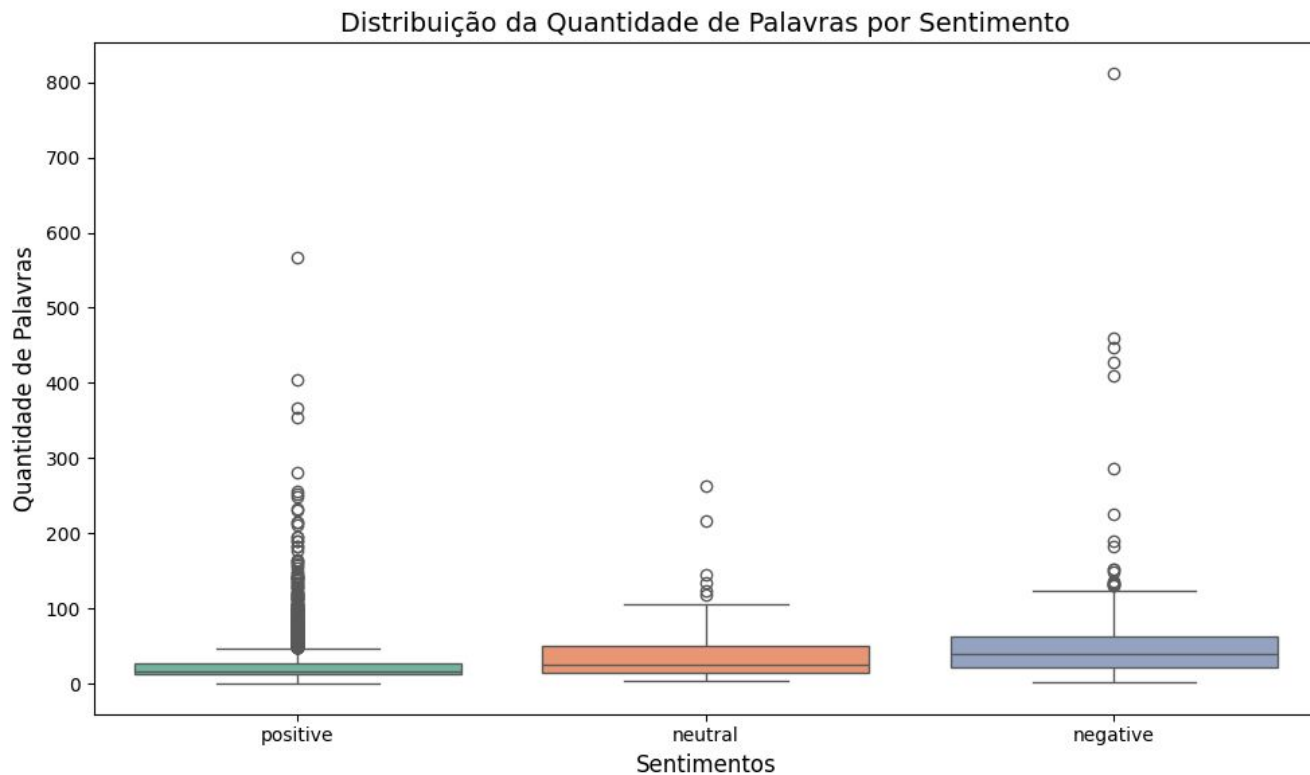
Nuvem de Palavras - Todos os Reviews



Análise Exploratória - Distribuição de sentimentos



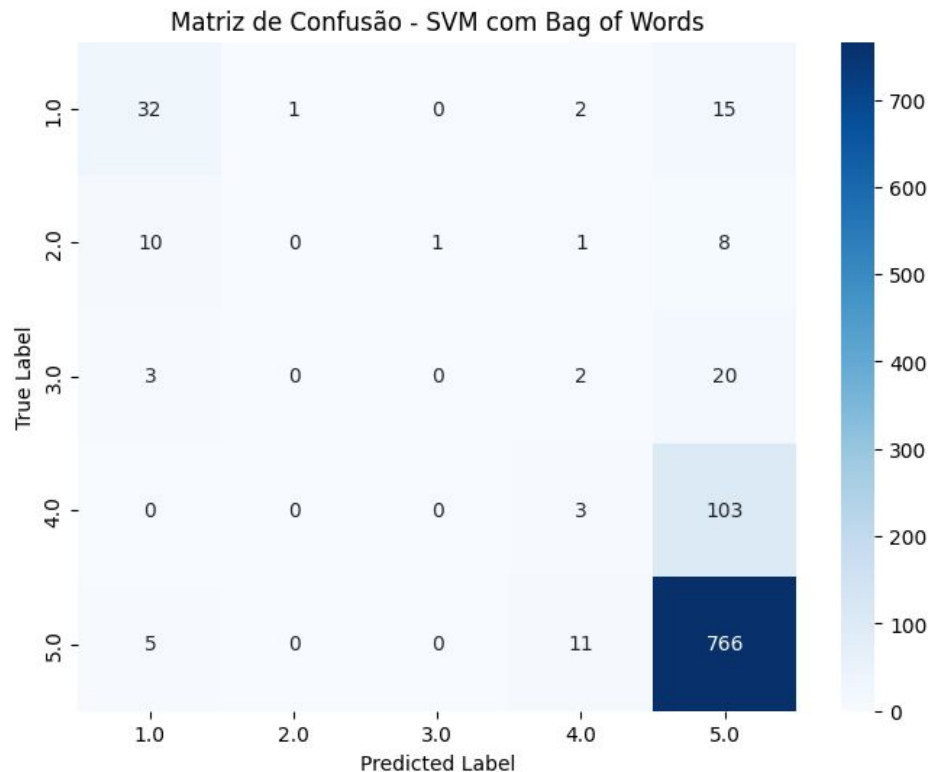
Análise Exploratória - Boxplot de palavras por sentimento



Modelagem SVM com bag of words

Bag of Words:

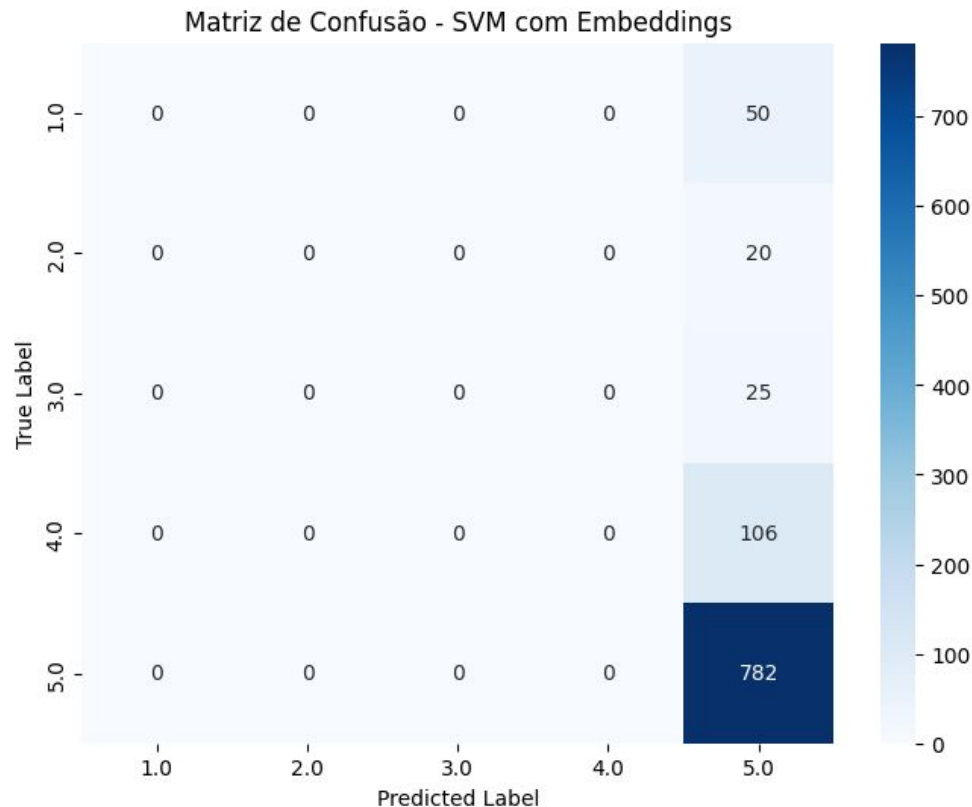
- Conversão para matrizes de frequência.
- Busca de hiperparâmetros com GridSearch.



Modelagem com SVM

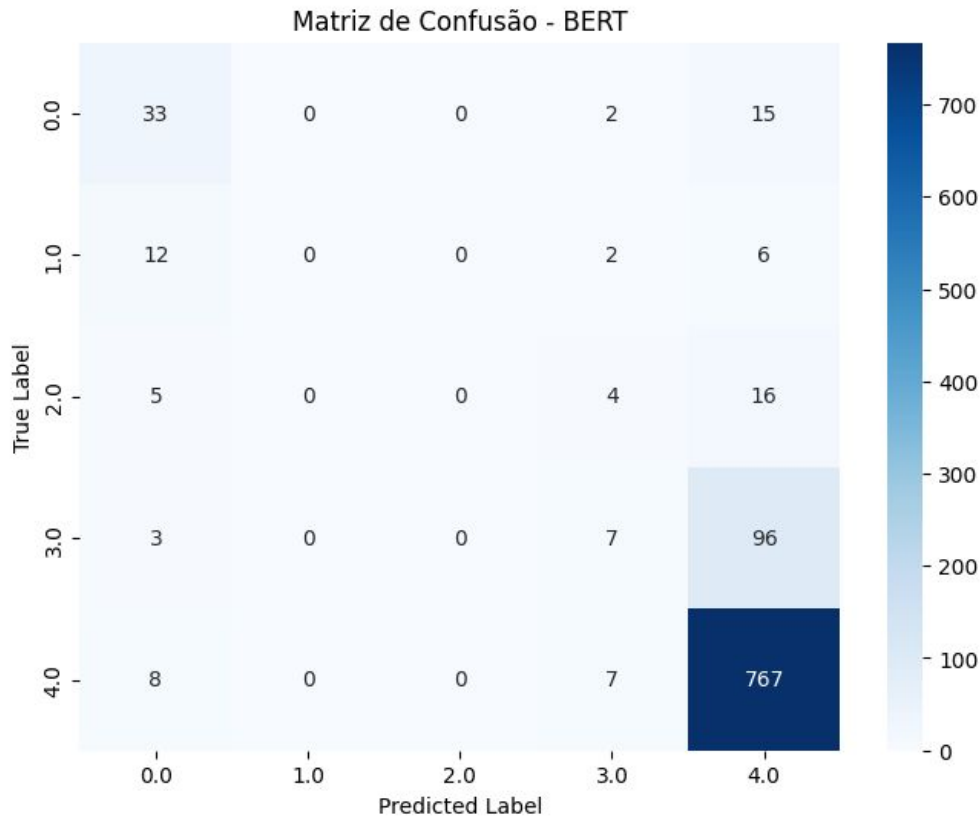
Embeddings (spaCy):

- Vetorização semântica dos textos.
- Normalização com StandardScaler.



Modelagem com BERT

- Uso de pré-treinamento.
- Fine-tuning no conjunto de reviews.
- Checkpoints para reaproveitamento.



Resultados

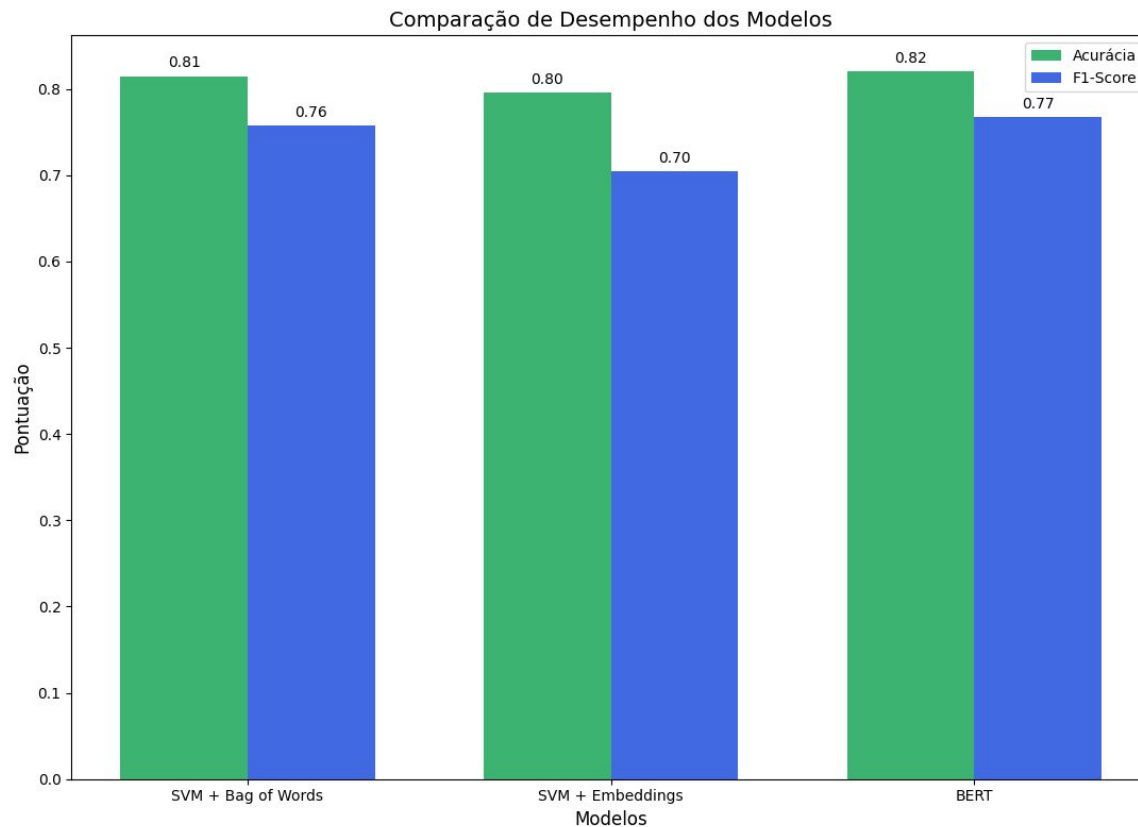


Gráfico: Comparação de desempenho (Acurácia e F1-Score).

Conclusão e Próximos Passos

Conclusão:

- BERT se destaca em tarefas complexas de NLP.
- SVM pode ser eficaz para abordagens leves.

Próximos Passos:

- Explorar mais dados.
- Testar técnicas de aumento de dados.
- Experimentar outros modelos de linguagem.

Obrigado!!

Pedro Gabriel Santos do Couto Soares | pgscs@cin.ufpe.br
Carlos Eduardo Figueiredo Costa | cefc@cin.ufpe.br

Especialização
Deep Learning

