

Classificação de Reviews de Produtos

Relatório de Análise e Machine Learning

Este relatório apresenta os resultados de um modelo de Machine Learning desenvolvido para classificar automaticamente reviews de produtos como positivos ou negativos.

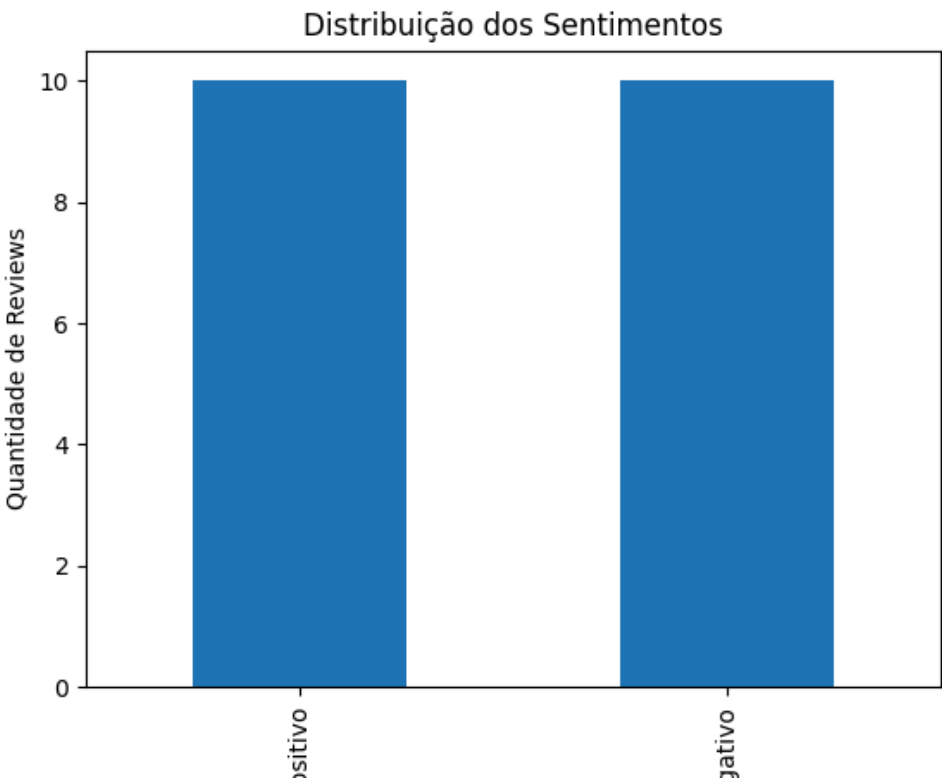
Os dados utilizados são fictícios e têm fins exclusivamente educacionais.

Contexto e Objetivo

A análise foi realizada com o objetivo de avaliar a viabilidade de automatizar a classificação de sentimentos em reviews de produtos, utilizando técnicas de Processamento de Linguagem Natural (NLP).

Visão Geral dos Dados

O conjunto de dados contém reviews textuais de produtos eletrônicos, rotulados como positivos ou negativos.



Metodologia

O processo analítico seguiu as seguintes etapas:

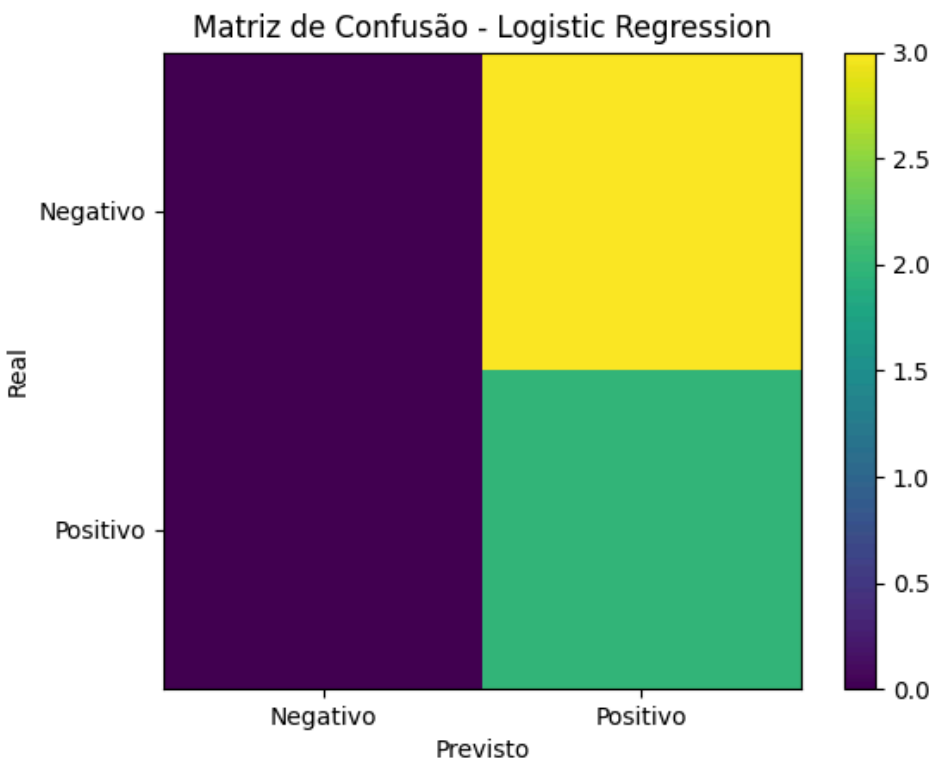
- Análise exploratória dos dados
- Pré-processamento dos textos
- Vetorização TF-IDF
- Treinamento de modelo de Regressão Logística
- Avaliação do desempenho

Resultados

O modelo apresentou desempenho limitado, o que era esperado dado o tamanho reduzido do dataset e a curta extensão dos textos analisados.

O modelo apresentou accuracy de 40%. No entanto, essa métrica deve ser interpretada com cautela devido ao tamanho reduzido do conjunto de teste, composto por apenas cinco observações.

Observa-se que o modelo apresentou viés em favor da classe positiva, classificando a maioria dos reviews como positivos, o que impactou negativamente sua capacidade de identificar avaliações negativas.



Apesar do desempenho limitado, o projeto cumpre seu objetivo ao demonstrar um pipeline completo de classificação de sentimentos, evidenciando desafios reais associados a dados textuais pequenos e ruidosos.

In []: