

# Classificação de Sentimento em Textos Utilizando Aprendizado de Máquina

Livia N. Cantowitz

Faculdade de Computação e Informática – Universidade Presbiteriana Mackenzie  
(UPM)

10389419@mackenzista.com.br

**Resumo.** *Este projeto tem como objetivo desenvolver um classificador de sentimentos capaz de identificar se um texto expressa opinião positiva ou negativa. O problema será abordado por meio de técnicas de Processamento de Linguagem Natural (PLN) e aprendizado supervisionado. O trabalho contempla a coleta de dados públicos, análise exploratória e preparação dos textos, além da implementação de algoritmos de classificação. Espera-se como resultado um modelo funcional e eficiente, capaz de diferenciar sentimentos com precisão satisfatória.*

## 1. Introdução

### a) Contextualização

A análise de sentimentos é uma das aplicações mais úteis do PLN, sendo utilizada em áreas como marketing digital, atendimento ao cliente e monitoramento de redes sociais. A partir de comentários, avaliações ou postagens, é possível identificar a percepção de usuários sobre produtos, serviços e eventos.

### b) Justificativa

Com o grande volume de textos disponíveis online, a classificação automática de sentimentos tornou-se essencial para transformar opiniões não estruturadas em informações úteis. O projeto é viável, educacional e pode ser realizado de forma simples, utilizando datasets disponíveis.

### c) Objetivo

Desenvolver um classificador de sentimentos que diferencie textos positivos e negativos utilizando aprendizado de máquina.

### d) Opção do Projeto

Opção escolhida: **ML/DL/VC/PLN**

## 2. Fundamentação Teórica

**Processamento de Linguagem Natural (PLN):** técnicas para representar e analisar dados textuais.

**Bag-of-Words / TF-IDF:** formas clássicas de transformar textos em vetores numéricos.

**Classificadores supervisionados:** algoritmos como Naive Bayes, Regressão Logística e SVM são amplamente usados em análise de sentimentos.

**Trabalhos relacionados:** estudos mostram que modelos simples já conseguem boa performance em datasets de sentimento.

### 3. Descrição do Problema

O problema consiste em classificar automaticamente opiniões em duas categorias: positiva ou negativa. Isso permite identificar rapidamente a percepção geral de usuários sem depender de leitura de textos.

### 4. Aspectos Éticos do uso da IA e Responsabilidade no desenvolvimento da solução

Embora útil, a análise de sentimentos pode gerar riscos quando usada sem transparência, como manipulação de dados de opinião ou decisões enviesadas. O projeto assume compromisso, garantindo que os dados utilizados sejam públicos e anonimizados, e destacando a importância de interpretação crítica dos resultados.

### 5. Dataset, se for o caso (anonimizados quando necessário), descrição detalhada do seu conteúdo/origem, análise exploratória e preparação dos dados em Python;

**Origem:** datasets públicos como *IMDb Reviews* ou *Twitter Sentiment Analysis* (Kaggle).

**Conteúdo:** textos curtos (frases/comentários) rotulados como “positivo” ou “negativo”.

**Anonimização:** os textos não conterão dados pessoais identificáveis.

#### Análise exploratória:

- Contagem de textos por classe.
- Exemplos de frases positivas e negativas.
- Distribuição do tamanho dos textos.

#### Preparação:

- Limpeza dos textos (remoção de stopwords, pontuação).
- Vetorização (TF-IDF).
- Divisão treino/teste.

## **6. Metodologia e Resultados Esperados**

### **Metodologia:**

1. Coleta do dataset público (IMDb ou Twitter Sentiment).
2. Pré-processamento dos textos (limpeza, tokenização, vetorização).
3. Treinamento de classificadores supervisionados.
4. Avaliação com métricas: acurácia, precisão, recall e F1-score.

### **Resultados esperados:**

1. Um modelo funcional para classificar sentimentos.
2. Relatório detalhado da análise exploratória e resultados.
3. Código público no GitHub.

## **Bibliografia**

Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow.

Bird, S., Klein, E., & Loper, E. (2009). Natural Language Processing with Python.

Russell, S., & Norvig, P. (2010). Artificial Intelligence: A Modern Approach.