



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL RURAL DA AMAZÔNIA
CAMPUS CAPITÃO POÇO
BACHARELADO EM SISTEMAS DE INFORMAÇÃO

INTRODUÇÃO À INTELIGÊNCIA ARTIFICIAL

Capitão poço – PA

2025

Relatório: Análise de Sentimentos de Avaliações de Filmes

1. Objetivo do Código

O projeto tem como objetivo aplicar técnicas de Processamento de Linguagem Natural (PLN) para identificar e classificar os sentimentos presentes nas avaliações de filmes fornecidas pelos usuários. A análise é realizada tanto em Python quanto em R, utilizando diferentes abordagens e bibliotecas.

2. Funcionamento do Código em Python

Entrada de Dados:

- O usuário faz upload de um arquivo CSV contendo duas colunas: filme e avaliacao.
- Processo:
- Cada avaliação é traduzida para o inglês utilizando a biblioteca googletrans, permitindo o uso do analisador de sentimentos SentimentIntensityAnalyzer do NLTK, que possui vocabulário baseado no idioma inglês.
- A análise classifica cada avaliação como positiva, negativa ou neutra, de acordo com a pontuação de sentimento composto (compound).

Saída:

- Exibição da tabela com os sentimentos associados a cada filme.
- Visualização gráfica da distribuição dos sentimentos (gráfico de barras).
- Opção para download dos resultados em arquivo CSV.

3. Funcionamento do Código em R

Entrada de Dados:

- Um vetor de textos contendo avaliações de usuários é pré-definido no código.

Processo:

- O texto é tokenizado em palavras utilizando unnest_tokens.

- É feito um `inner_join` com um dicionário de sentimentos em português contendo palavras positivas e negativas.

Com base na contagem de palavras positivas e negativas encontradas em cada texto, o sentimento final é classificado.

Saída:

- Exibição dos sentimentos detectados para cada texto.
- Geração de um gráfico de barras com a contagem de avaliações positivas, negativas e neutras.
- O gráfico é salvo automaticamente como `grafico_sentimentos.png`.

4. Comparação entre as Abordagens

Critério	Python (NLTK + Streamlit)	R (tidytext + ggplot2)
Idioma da Análise	Traduz as avaliações para o inglês antes de analisar	Análise direta em português
Tipo de Modelo	Modelo pré-treinado VADER (lexical e estatístico)	Dicionário manual com palavras positivas e negativas
Interface	Web interativa via Streamlit	Script executado no RStudio ou terminal
Conexão com a Internet	Necessária para traduzir os textos com o Google Translate	Funciona completamente offline
Facilidade de Uso	Muito fácil para usuários leigos, com interface gráfica amigável	Requer conhecimentos básicos de R e manipulação de dados
Visualização	Gráfico de barras interativo + opção de download do CSV	Gráfico estático gerado com `ggplot2` e salvo como imagem
Precisão Linguística	Pode perder nuances durante a tradução	Preserva o contexto original, mas depende da qualidade do dicionário
Customização	Limitada ao funcionamento do NLTK/VADER	Alta, podendo expandir o dicionário conforme o contexto

Resultado:

