



## ***Deep Learning Para Classificação de Sentimento em Vídeos utilizando legendas***

**Jean Ferreira DIAS<sup>1</sup>; Diego SAQUI<sup>2</sup>; Heber Rocha MOREIRA<sup>3</sup>.**

### **RESUMO**

Este projeto de pesquisa explora o uso de técnicas de Deep Learning aplicadas à análise de sentimentos em vídeos do Youtube, utilizando legendas geradas automaticamente como fonte de dados. A análise de sentimentos permite a classificação de emoções expressas durante os vídeos, gerando dados que tornam possível analisar a correlação das mesmas com métricas de engajamento, como visualizações e interações. A proposta envolve o desenvolvimento de uma base de dados com legendas de vídeos do YouTube e um modelo de aprendizado profundo, permitindo uma avaliação mais precisa da relação entre os sentimentos e o desempenho do conteúdo audiovisual. Os resultados esperados incluem a identificação de padrões emocionais em vídeos populares e insights para criadores de conteúdo e profissionais de marketing.

**Palavras-chave:** Inteligência Artificial, Processamento de Linguagem Natural, Engajamento, Redes Sociais.

### **1. INTRODUÇÃO**

Nos últimos anos, o avanço dos dispositivos móveis e a ascensão das redes sociais transformaram a forma como as pessoas consomem conteúdo audiovisual. Segundo dados do Google (2021), aproximadamente 113 milhões de brasileiros consomem vídeos no YouTube mensalmente, o que demonstra a importância desse formato de conteúdo.

A criação de vídeos, no entanto, apresenta desafios. Um dos principais é a dificuldade de obter dados significativos a partir dos vídeos, que possam ajudar criadores e empresas a entenderem melhor o comportamento do público e a planejar novas estratégias de engajamento. Pereira (2018) argumenta que o uso de técnicas computacionais avançadas, como o Processamento de Linguagem Natural (PLN), pode facilitar essa análise, permitindo a interpretação automática das emoções expressas nos vídeos através das legendas.

Neste contexto, a análise de sentimentos é uma técnica poderosa, que utiliza o PLN para identificar, extrair e classificar a polaridade emocional de um texto, seja ele positivo, neutro ou negativo (SAMPAIO, 2021). Este projeto visa aplicar técnicas de Deep Learning para realizar a análise de sentimentos em legendas de vídeos, permitindo assim a correlação entre os sentimentos expressos e os níveis de engajamento dos vídeos nas redes sociais.

<sup>1</sup>Jean Ferreira Dias, Bacharelado em Ciência da Computação, IFSULDEMINAS – Campus Muzambinho. E-mail: [jean.dias@alunos.ifsuldeminas.edu.br](mailto:jean.dias@alunos.ifsuldeminas.edu.br).

<sup>2</sup>Diego Saqui, Instituto de Ciência, Tecnologia e Inovação (ICTIN) - Universidade Federal de Lavras (UFLA) - Campus São Sebastião do Paraíso. E-mail: [diego.saqui@ufla.br](mailto:diego.saqui@ufla.br).

<sup>3</sup>Heber Rocha Moreira, IFSULDEMINAS – Campus Muzambinho. E-mail: [heber.moreira@ifsuldeminas.edu.br](mailto:heber.moreira@ifsuldeminas.edu.br).

## 2. FUNDAMENTAÇÃO TEÓRICA

Para compreender as bases deste projeto, é necessário entender alguns conceitos fundamentais, que serão apresentados neste tópico.

Começando pela Inteligência Artificial (IA), que é a área da ciência da computação que se dedica à criação de sistemas capazes de realizar tarefas que normalmente requerem a inteligência humana. Cozman, Plonski e Neri (2021) definem a IA como o campo que busca construir artefatos capazes de tomar decisões e resolver problemas de maneira autônoma. De forma geral, são algoritmos que analisam dados e tomam decisões tal qual um ser humano faria.

O Processamento de Linguagem Natural (PLN) é uma subárea da IA voltada para a compreensão e interpretação da linguagem humana por máquinas. Vieira e Lopes (2010) apontam que o PLN enfrenta o desafio das ambiguidades e nuances das linguagens naturais, exigindo técnicas específicas para diferentes contextos, como a análise de textos informais ou formais.

A Análise de Sentimentos é uma aplicação do PLN que busca identificar emoções em textos, classificando-os em categorias como positivo, negativo ou neutro. A análise pode ser feita através de abordagens léxicas, que utilizam dicionários de palavras já classificadas, ou de modelos de aprendizado de máquina, que identificam padrões nos textos.

Aprendizado de Máquina (*Machine Learning*) é o “campo da ciência que fornece ao computador a habilidade de aprender uma determinada tarefa sem ser explicitamente programada” (HOMEM, 2020). É possível realizar esse processo analisando conjuntos de dados robustos e encontrando padrões que relacionam esses dados. No caso da análise de sentimentos, o Aprendizado de Máquina busca relacionar os tokens com os sentimentos expressados nos textos, encontrando palavras específicas que expressam sentimentos específicos.

O Aprendizado Profundo (*Deep Learning*) é uma vertente do aprendizado de máquina que utiliza redes neurais artificiais para realizar tarefas complexas, como o reconhecimento de padrões em grandes volumes de dados. Chaves (2021) destaca que as arquiteturas de redes neurais recorrentes (RNN) e convolucionais (CNN) têm se mostrado especialmente eficazes em tarefas de processamento de linguagem natural e análise de sentimentos.

## 3. MATERIAL E MÉTODOS

Para alcançar os objetivos propostos nesta pesquisa, foram seguidos alguns passos metodológicos, que incluem desde a coleta até a análise dos dados. Inicialmente uma base de dados foi construída a partir do dataset “Youtube Trending Video Dataset”, disponível na plataforma Kaggle. A base de dados foi complementada com as legendas dos vídeos em língua portuguesa utilizando a API do YouTube e bibliotecas Python. As legendas passaram por um pré

processamento, para garantir a qualidade dos dados, que incluiu a remoção de stopwords (palavras sem relevância significativa) e tokenização (segmentação do texto em palavras individuais).

A partir dessa base de dados foram empregadas duas abordagens: uma automática e outra manual. Na abordagem automática, foi utilizado o algoritmo *Valence Aware Dictionary for Sentiment Reasoning* (VADER), que atribui *scores* de sentimentos a cada frase. Na abordagem manual, uma amostra das legendas foi classificada individualmente, atribuindo-se polaridades emocionais com base na interpretação do contexto.

A partir da biblioteca Keras do Python foi implementado um modelo sequencial de Deep Learning, contendo duas camadas ocultas e uma camada de saída para classificação binária, com o propósito de analisar ambas as bases de dados resultantes das classificações automática e manual. Os dados foram divididos em conjuntos de treinamento e teste, na proporção de 80% para treinamento e 20% para teste.

#### 4. RESULTADOS E DISCUSSÃO

Para avaliar o modelo de análise de sentimentos desenvolvido, foram realizados dois experimentos. No primeiro, a rede neural foi treinada com a base de dados feita com o classificador VADER. O modelo obteve uma acurácia de 0.88, o que indica um bom desempenho na identificação de emoções em legendas de vídeos.

No segundo teste, o modelo foi treinado e testado utilizando classificações manuais. Esse método resultou em uma acurácia de 0.78, inferior ao primeiro teste. A diferença entre esses testes pode ser explicada pela complexidade de se simular a forma que humanos analisam frases. Ao contrário do algoritmo VADER, que classifica através das pontuações individuais de cada palavra, humanos levam em consideração todo o contexto da frase, muitas vezes envolvendo ambiguidades e sutilezas.

A comparação dos dois testes sugere que, embora o VADER tenha tido uma melhor performance em termos de acurácia, a rotulação manual pode, em contextos específicos, fornecer uma visão mais profunda e diferenciada dos sentimentos. No entanto, a subjetividade humana, quando combinada com a necessidade de modelos treinados para interpretar tais variações, ainda apresenta desafios. O modelo ainda tem que ser refinado para conseguir uma acurácia melhor nesse contexto específico.

Tipo de Classificação	Acurácia
Classificador Vader	88%
Classificações Manuais	78%

Fonte: do autor

## 5. CONCLUSÃO

Este projeto demonstrou a viabilidade de aplicar técnicas de Deep Learning e Processamento de Linguagem Natural para a análise de sentimentos em legendas de vídeos do YouTube. Através do desenvolvimento de uma base de dados e de um modelo sequencial de redes neurais, foram realizados testes comparando classificações automáticas e manuais de sentimentos.

Os resultados mostram que, embora o classificador VADER tenha maior acurácia, a classificação manual captura nuances emocionais desafiadoras para os modelos de deep learning. A diferença entre os testes evidencia a dificuldade dos modelos em replicar completamente a subjetividade humana. O estudo destaca a necessidade de continuar refinando as redes neurais e explorar soluções híbridas que combinam classificadores automáticos e manuais para melhorar a interpretação de sentimentos.

## REFERÊNCIAS

CHAVES, Leticia Saraiva. Utilizando um modelo transformer no processo de identificação de entidades nomeadas em textos criminais. 2021.

COZMAN, Fabio G.; PLONSKI, Guilherme Ary; NERI, Hugo. Inteligência Artificial. 2021.

HOMEM, William Ludovico; UFES, P. E. M. Apostila de machine learning. PET Engenharia Mecânica, UFES, 2020.

PEREIRA, Moisés Henrique Ramos. Análise de sentimentos Multimodal Aplicada à Computação de Níveis de Tensão em Vídeos de Notícias. Tese de Doutorado. Centro Federal de Educação Tecnológica de Minas Gerais, 2018.

Pesquisa inédita mostra o impacto econômico, cultural e social do YouTube no Brasil, Think with Google Brasil, 2021. Disponível em:

<<https://www.thinkwithgoogle.com/intl/pt-br/estrategias-de-marketing/video/impacto-economico-cultural-social-youtube-brasil/18/03/2024>>. Acesso em: 09 de setembro de 2024.

SAMPAIO, Amanda Gomes. Análise de Sentimentos. Escola de Ciências Exatas e da Computação, 2021

VIEIRA, Renata; LOPES, Lucelene. Processamento de linguagem natural e o tratamento computacional de linguagens científicas. Em corpora, 2010.