

# UMA CLASSIFICAÇÃO DE MÚSICAS EM GÊNEROS VIA AS LETRAS

Ana Luisa Mendes  
analuisamd@ufmg.br

Maria Clara Ferreira  
mcaf@ufmg.br

## ABSTRACT

Lyrics stand out as one of the most crucial features that differentiate songs from instrumental music. This project proposes the use of a Convolutional Neural Network (CNN) to classify songs by genre based on their lyrics. The approach combines natural language processing and convolutional neural networks to explore the relationships between song lyrics and their genres. This meta-article aims to discuss the project's solutions and explain its results.

## 1. INTRODUÇÃO

A música é uma forma de expressão artística que transcende as fronteiras culturais e emocionais e é uma parte fundamental da experiência humana. Na música, as letras desempenham um papel único ao adicionar camadas de significado e emoção à obra.

Este trabalho é motivado pela necessidade de compreender e classificar a música com base em suas letras, reconhecendo que as letras cantadas desempenham um papel crucial na definição do caráter e do contexto de uma música. Nesse sentido, propomos aplicar redes neurais convolucionais (CNN) para classificar músicas por gênero e explorar as ricas interações entre letras e diferentes estilos musicais.

## 2. DADOS E PRÉ-PROCESSAMENTO

Foi usado um dataset com músicas (letras e gênero) da Taylor Swift. Ele contém 560 músicas, divididas em 40 álbuns, sendo alguns regravações de trabalhos antigos da cantora. Cada álbum foi atribuído a um gênero pré-definido a fim de facilitar as análises e execução do algoritmo.

Como tratamento de dados, as letras foram tokenizadas, padronizadas, e foram removidos delas *stopwords* de inglês, pontuação, caracteres especiais, etc. Cada álbum foi mapeado para um gênero musical, no qual foi mapeado para outra coluna com one-hot encoding.

Os dados foram divididos em um conjunto de treino e outro de teste.

### 2.1 Exploração dos dados

Foi feita uma análise de sentimentos com as letras das músicas, utilizando a biblioteca *TextBlob*, em que -1 significa um sentimento negativo e 1 positivo.

Pode ser visto nas figuras 1 e 2 boxplots da diversidade léxica de cada gênero (real), e da análise de sentimentos. Na figura 4 temos um plot da polaridade dos sentimentos contra a diversidade lexical.

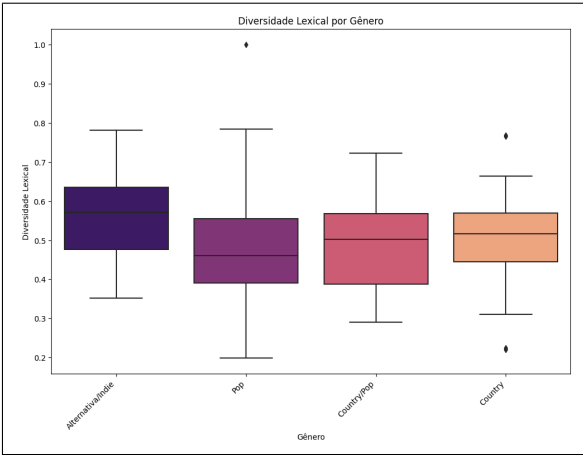


Figure 1. Diversidade lexical em cada gênero

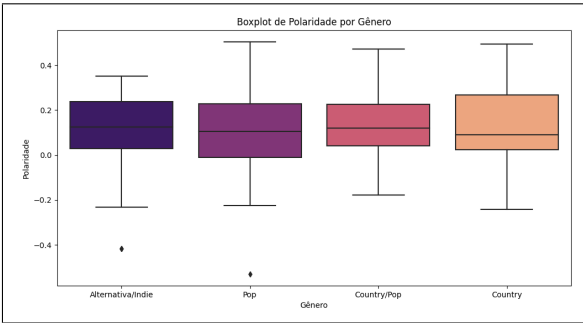


Figure 2. Polaridade de sentimentos por gênero

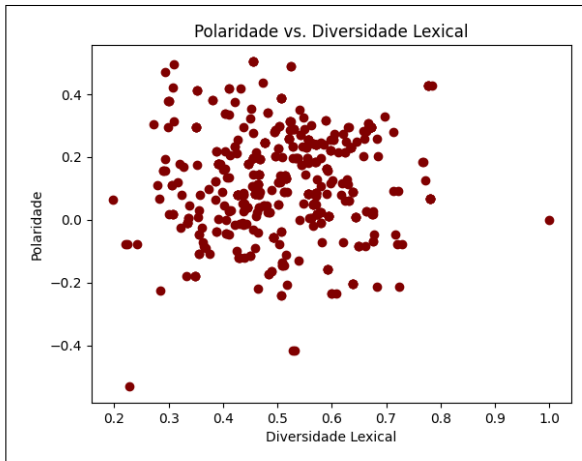
Como pode ser visto nas figuras 1 e 3, a diversidade lexical para o gênero alternativo é maior do que para os outros.

## 3. MODELO

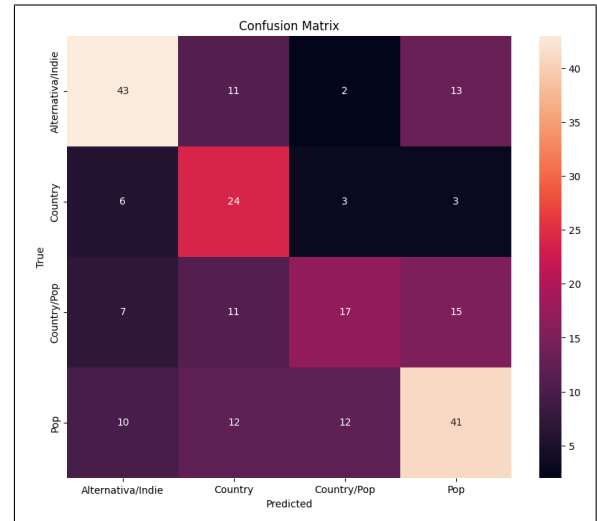
A arquitetura de um modelo sequencial com uma Rede Neural Convolutacional (CNN) é projetada para processar dados sequenciais, como sequências de texto. A camada de incorporação mapeia palavras para vetores densos, a camada convolutacional unidimensional detecta padrões locais importantes, e a camada de pooling global reduz a dimensionalidade preservando características cruciais.



Figure 3. Wordcloud para cada gênero



**Figure 4.** Diversidade lexical x polaridade de sentimentos



**Figure 5.** Matriz de confusão

A utilização de camadas densas, como Dense, permite aprender padrões mais complexos, enquanto a camada de desativação ajuda a evitar o overfitting. Essa arquitetura é eficaz para processar linguagem natural, pois captura padrões locais, reduz a dimensionalidade e é invariante a translações, tornando-se útil para tarefas como classificação de texto e por isso foi escolhida para a realização do projeto. O compartilhamento de parâmetros nas camadas convolucionais aprimora a capacidade do modelo de reconhecer padrões relevantes em diferentes partes da sequência o que faz com que a classificação dos gêneros musicais tenha uma grande taxa de acerto.

Foi criado um modelo sequencial com uma rede neural convolucional, usando TensorFlow. O modelo tem 3 camadas:

- camada de embedding de palavras, em que cada palavra é convertida em um vetor
- camada de LSTM, que é uma rede neural recorrente, que captura dependências de longo termo nos dados sequenciais
- camada densa de saída, que faz as previsões de gênero

O modelo foi então treinado e validado com os dados de treino

## 4. RESULTADOS

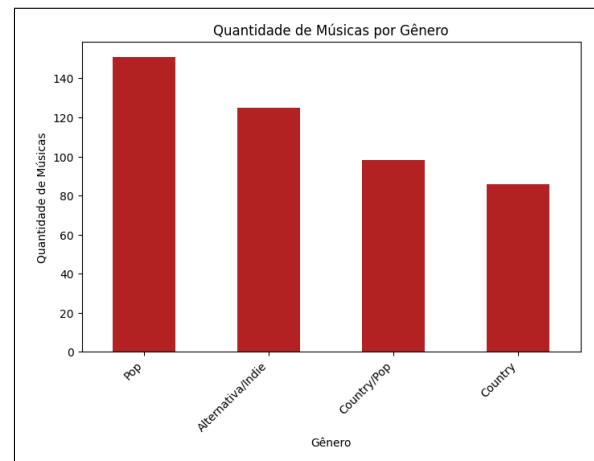
A acurácia do modelo nos dados de teste foi de 0.54.

### 4.1 Quantidade de músicas por gênero

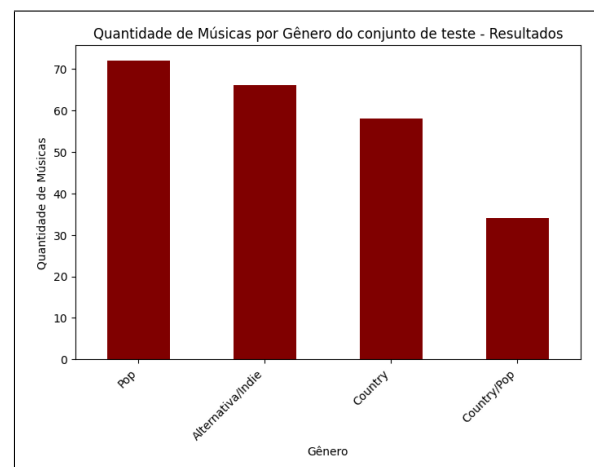
A quantidade de músicas por cada gênero permaneceu proporcional nos dados reais quanto nas previsões feitas pelo modelo.

## 5. CONCLUSÃO

Ao longo deste estudo, exploramos a aplicação de uma Convolutional Neural Network (CNN) para a classificação



**Figure 6.** Em todos os dados reais



**Figure 7.** No conjunto de teste

88 de músicas com base em suas letras, enfatizando a influên-  
89 cia das palavras na definição dos gêneros musicais. Os re-  
90 sultados obtidos mostram eficácia da abordagem proposta,  
91 destacando a capacidade da CNN em capturar padrões com-  
92 plexos e representações significativas nas sequências de  
93 texto das letras musicais.

94 A análise dos resultados demonstrou não apenas a habil-  
95 idade do modelo em distinguir gêneros musicais mas tam-  
96 bém ofereceu análises sobre as características linguísticas  
97 que contribuem para essa diferenciação. A aplicação da  
98 CNN proporcionou uma compreensão mais profunda das  
99 relações semânticas e estilísticas presentes nas letras, am-  
100 pliando nosso entendimento sobre como a linguagem in-  
101 fluencia a categorização musical.

## 102 6. REFERENCES

103 [1]Meinard Müller and Frank Zalkow: FMP Notebooks:  
104 Educational Material for Teaching and Learning Funda-  
105 mentals of Music Processing. Proceedings of the Inter-  
106 national Conference on Music Information Retrieval (IS-  
107 MIR), Delft, The Netherlands, 2019.

108 [2]Akshi Kumar, Arjun Rajpal, and Dushyant Rathore.  
109 Genre classification using word embeddings and deep learn-  
110 ing. 2018 International Conference on Advances in Com-  
111 puting, Communications and Informatics (ICACCI), 2018.

112 [3] Arjun Raj Rajanna, Kamelia Aryafar, Ali Shoko-  
113 ufandeh, and Raymond Ptucha. Deep neural networks: A  
114 case study for music genre classification. 2015 IEEE 14th  
115 International Conference on Machine Learning and Appli-  
116 cations (ICMLA), 2015.