Leveraging Song Lyrics to Predict Music Genres

Motivation:

Genre classification plays an important role in the music industry, influencing playlist curation, music recommendation algorithms, and listener engagement. Accurately identifying a song's genre can improve music streaming platforms, allowing them to deliver more personalized content. Also, understanding how different genres are defined through lyrical content could better our understanding of music trends and cultural influences.

This project aims to build a model that classifies songs based on their lyrics by analyzing how vernaculars differ across genres. For instance, country songs might use specific themes or phrases that set them apart from rock or pop. Identifying these patterns could lead to a better understanding of differences in genres or even the creation of new sub-genres. This work could impact on how we define and categorize music in an industry where genres are becoming more fluid.

Deliverable:

Your deliverable is to develop a text-based Machine Learning model that can accurately classify music genres at up to 60% accuracy, to set a benchmark model that can be eventually expanded upon. Deliver a report summarizing the steps you took in selecting your model and its parameters and why you chose to do so. Additionally, you must perform exploratory data analysis on your data and find actionable insights that can be used to help educate our understanding of genres.

GitHub: https://github.com/mthabet1/CS3