Analyzing Trends in Emotional Tone of Song Lyrics and Their Relationship with Mental Health Disorders

Ömer Yıldırım, Jianwen Cao, Yuxuan Wang

September 19, 2024

1 Project Summary

We will try to analyze the evolution of song lyrics over time to detect trends in emotional tone, including elements like aggressive or harmful language and themes of connection or love, and examine the correlation between these trends and global mental health disorder patterns.

2 Project Overview

This project involves the analysis of song lyrics across time to detect trends in emotional tone and language using sentiment analysis and hate speech detection. The goal is to examine how lyrical content may have shifted toward more aggressive or negative expressions over the years and to explore the potential correlation between these shifts and global mental health disorder trends.

Since music is a powerful medium that reflects societal trends, exploring how it has evolved in relation to mental health could provide valuable insights into its influence on public well-being. This project also allows us to apply NLP techniques to a real-world problem, providing an opportunity for meaningful research that could yield novel insights.

Current research on music and mental health primarily focuses on therapeutic effects. The project aims to extend this by exploring music as a potential contributing factor to mental health disorders.

The datasets given below which can be accessed via Kaggle will provide the necessary data for analyzing the evolution of lyrical content in terms of their tone and examining their relationship with global mental health trends.

- 160k Spotify songs from 1921 to 2020 (Sorted) [1] : 160,000+ tracks sorted by name, from 1921-2020 found in Spotify as of June 2020.
- 960K Spotify Songs With Lyrics [2]: Song attributes and lyrics with timestamps for 960K Spotify songs.
- Global Trends in Mental Health Disorder [3]: This dataset contains informative data from countries across the globe about the prevalence of mental health disorders.

We will build a sentiment analysis model to analyze the emotional tone of song lyrics over the decades. Using this model, we will generate sentiment scores calculated yearly to evaluate trends. We will also create a statistical analysis framework to explore the connection between these trends and mental health disorder data. The output will also include visualizations of trends, statistical analyses, and a detailed project report to construct a cohesive research.

Open questions for this project include whether the data will reveal a clear link between negative lyrical trends and mental health disorders, and if so, how strong that relationship will be.

References

[1] *160k Spotify songs from 1921 to 2020 (Sorted)*. en. URL: https://www.kaggle.com/datasets/fcpercival/160k-spotify-songs-sorted (visited on 09/19/2024).

[2] *960K Spotify Songs With Lyrics data*. en. URL: https://www.kaggle.com/datasets/bwandowando/spotify-songs-with-attributes-and-lyrics (visited on 09/19/2024).

[3] *Global Trends in Mental Health Disorder*. en. URL: https://www.kaggle.com/datasets/thedevastator/uncover-global-trends-in-mental-health-disorder (visited on 09/19/2024).