Emotion Detection on song lyrics stanzas

TEXT ANALYTICS

Group 2

Barbieri, Bosco, Ferrara, Rotellini, Zizza

Index

Elenco delle figure

Introduction

Team members and roles

Motivation and project goal

Songs have the unique ability to engage people in ways that few other mediums can match. While beats and melodies contribute to this impact, it is often the lyrics that give songs their true emotional strength. As a matter of fact, lyrics serve as the foundation of songs, playing the crucial role of expressing complex human emotions and moods in many different ways.

Despite this, the general tendency is to focus more on song genres, themes, or artists, rather than on the specific emotions conveyed through lyrics. But by analyzing the emotional tone of song texts, people could actually gain valuable insights into cultural trends, social issues, and individual mental states. Furthermore, this type of analysis could help identify how specific emotional responses are evoked, making it a useful tool across various fields.

Therefore, the main goal of this project would be to perform an emotion detection task of each stanza of the songs in the selected dataset, so as to gain a deeper understanding of the fluctuations of the emotions inside the songs. Furthermore, we intend to compare different textual preprocessing and Machine Learning models, in order to explore some different possibilities and evaluate their performances.

1. Dataset overview

The dataset used for the project is a set of lyrics of songs extracted from the Genius Song Lyrics Dataset ([1]). The dataset contains 11 attributes that represent various song data, including the lyrics. It has a total of over 3 million songs by 641,349 different artists. The original dataset includes songs in all languages: for our aim we will be using the english ones only. For now we are keeping every other column including, of course, the lyrics one, since we might be able to use them. Then, each song will be divided into its composing stanzas, which are the actual subject of our analysis.

In order to create the ground truth to train the models, we will be using the already existing Meta's Bart large Multi Natural Language Inference (MNLI). We will be using Robert Plutchik's 8 primary emotions as labels; the emotions included in this representation are: anger, fear, sadness, disgust, surprise, anticipation, trust, and joy. Such multifaceted emotions allow us to finely analyze the feelings and moods conveyed by songs.

2. Preprocessing

3. Static Models

4. Neural Networks

Key findings and conclusions

Bibliografia

[1] *Genius Song Lyrics*. URL: https://www.kaggle.com/datasets/carlosgdcj/genius-song-lyrics-with-language-information?select=song_lyrics.csv.