Emotion Detection on song lyrics stanzas

TEXT ANALYTICS

Group 2

Barbieri, Bosco, Ferrara, Rotellini, Zizza

Index

Elenco delle figure

Introduction

This report illustrates development and findings of Group 2's project for the course Text Analytics, Academic Year 2024-2025.

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. Lyrics serve as one of the main foundations of songs, playing a crucial role in expressing feelings in many different ways. Analyzing the emotional tone of song texts can give useful insights about individual mental states, cultural trends, social issues and more.
This emotional tone data can be useful for many things, such as automatized playlist creation, or songs' organization, offering an alternative to the more traditional genre-based classification.

The project's goal was to develop various Machine Learning models that perform this task. Instead of detecting emotions for entire songs, the models assign emotion labels to individual stanzas, to obtain a deeper understanding of emotional fluctuations within the lyrics. The emotional classification is based on Robert Plutchik's eight primary emotions (shown in figure 1), offering a comprehensive range for representing diverse emotional states.

Figura 1: Plutchik's eight primary emotions

Furthermore, we intend to compare different textual preprocessing and Machine Learning models, in order to explore different possibilities and evaluate their performances.

Chapter overview

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. The original dataset includes songs in all languages: for our aim we will be using the english ones only.

1.1 Ground Truth

The dataset has no emotion label, which is needed for training the models. To create the ground truth for models' training, the model Albert Base v2[2] was used. The stanzas are labeled using Robert Plutchik's 8 primary emotions; 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.

[2] *Albert Base v2*. URL: https://huggingface.co/albert/albert-base-v2.