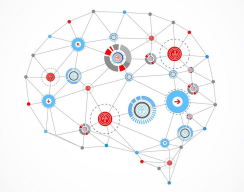
Mood Classification of Song Lyrics Using Deep Learning

Cyprian Gascoigne, Jack Workman, Yuchen Zhang
Instructor: Daniel Cer

Abstract



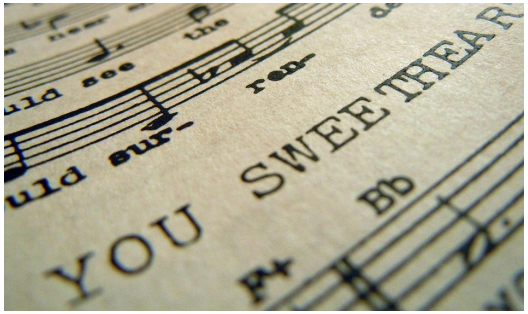
Mood classification of lyrics can help in the creation of automatic playlists, a music search engine, labeling for digital music libraries, and other recommendation systems.

We compare the accuracy of deep learning vs traditional machine learning approaches for classifying the mood of songs, using lyrics as features.

We will use mood categories derived from Russell's model of affect (from psychology, where mood is represented by vector in 2D valence-arousal space) and also calculate a valence-happiness rating.

We will test the extensibility of our deep learning model using the Million Song Dataset (MSD), a freely available contemporary music track dataset.

Introduction



Music recommendation engines are becoming more and more expected norms for listeners and music lovers today. Since mood and emotional classification is crucial to the entertainment value of songs, music recommendation algorithms can take into account mood classification.

Song lyrics are different from ordinary text in that they often use more stylistic qualities like rhyming, and other forms, often contributing to the emotional value (Lee et al., 2010). They are also shorter and often have smaller vocabularies than other text documents. This, combined with their more poetic style can cause more ambiguity when it comes to mood identification (Cano et al., 2017).

Methodology

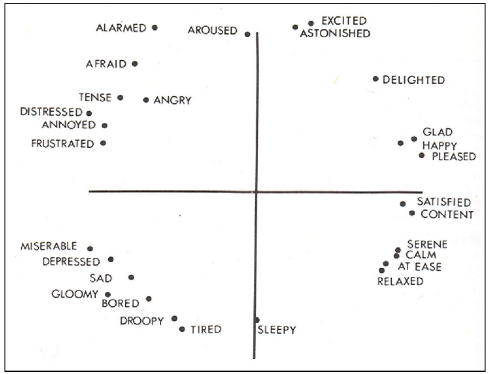
Data Acquisition

One initial challenge was in acquiring corresponding lyrics to song data. To obtain the 35,000+ lyrics in our dataset, we used the python package lyricsgenius for retrieving lyrics. The package interfaces with the www.genius.com API for lyric access, and iterates through the artist-song pairs in the Musixmatch file.

Word Embeddings

To form our word embeddings, we make use of the word2vec model (Mikolov et al., 2013) and the implementation provided by TensorFlow.

Methodology (cont.)



Mood Annotation

We use the popular Russell's model of affect and focus on 18 mood classes similar to previous studies (Chen et al., 2009). In Russell's model, mood is represented by a vector in the 2-D valence-arousal space where mood is an element of (valence, arousal): m M = (v, a)

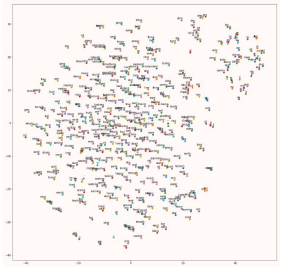
Mapping Labels

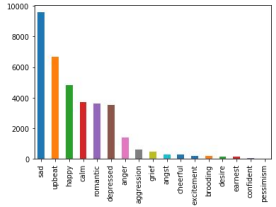
Once we have the index built, we can easily match the lyrics to the mood tags from the Last.fm dataset. From MSD, we use the Last.fm and musiXmatch datasets for song tags and lyrics.

Methodology (cont.)

Model Training

The remaining lyrics are then randomly divided into a training set, dev set, and test set, each with similar mood distributions (21,501; 7,167; 7,167 lyrics respectively). Then we establish a baseline classification using simple classifiers. Finally, we can use our neural network architecture (CNN) for modeling mood classification.





Results

Conclusion