

# 605.744: Information Retrieval

## Project Proposal: Emotion Analysis of Song Lyrics

Sabbir Ahmed

October 17, 2022

The DepecheMood Emotion Lexica (<https://arxiv.org/abs/1405.1605> and <https://arxiv.org/abs/1810.03660>) provide frequencies of words per emotions extracted from crowd-annotated news. The emotions include: afraid, amused, angry, annoyed, dont\_care, happy, inspired, and sad.

The vocabulary of the lexica are relatively limited; less than 175,000 unstemmed terms across the datasets. However, the vocabulary consist of common words in English, including those considered to be stopwords. These types of words are common in popular song lyrics as well. These lyrics can be easily gathered from a combination of datasets on Kaggle, the musiXmatch dataset (<http://millionsongdataset.com/musixmatch/>) or through web scraping. Extracting emotion annotations from song lyrics can help classify songs into mood-based playlists, similar to algorithms by popular music streaming services like Spotify.