Predicting Music Emotion with Social Media Discourse

Funding and Support:
URSA Engage – Oregon State
Layman Fellowship – Oregon State

By – Aidan Beery
Faculty Advisor – Dr. Patrick Donnelly

Oregon State University
COMPUTER SCIENCE

Abstract

Determining the emotional affect of a song is a challenging task in the music information retrieval domain. Our research seeks to contribute a framework for predicting musical affect based on commentary from social media platforms. We extract features from comments by analyzing the sentiment of individual words.

Introduction

Evaluating a song's affective qualities is a time consuming, expensive process which relies on human annotators to manually rate samples of music. If intelligent models were able to predict affect, it could advance research in music information retrieval. Our hypothesis is that social media conversations about specific songs contain semantic information which may be used to estimate song emotion.

Dataset

Our experiment relies on sentiment wordlists for feature generation and music emotive datasets for model training. We assess the AMG1608 (n=1608) list of songs, with human annotated valence (measure of positivity) and arousal (measure of energy)[1]. Our sentiment wordlists provide annotated ratings of various emotion metrics for single words.

Dataset	Number of words
NRC-EmoAff[2]	4986
NRC-EmoLex	14182
NRC-EmoVAD	20007
MPQA[3]	8221
Extended-ANEW[4]	13915

Methods

For each of the 1608 songs, we query Reddit for ten posts matching that song's title and artist name. We perform sentiment analysis on these comments, creating features by extracting words matching our semantic wordlists. By finding the average and variance in valence, arousal, dominance, and sentiments, we calculate the mean affect across comments for a given song.

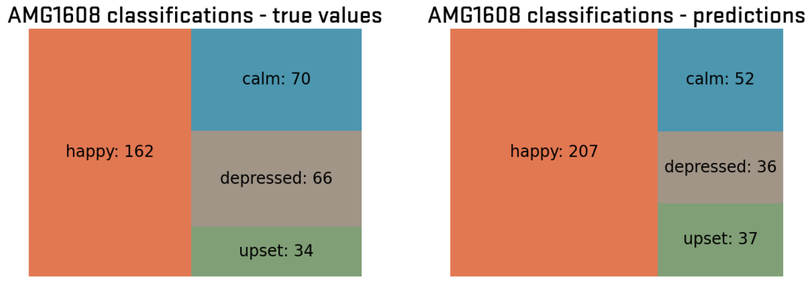

Figure 2 – A comparison of the predicted mood categories compared to the actual categories of the test set for an experiment.

Results

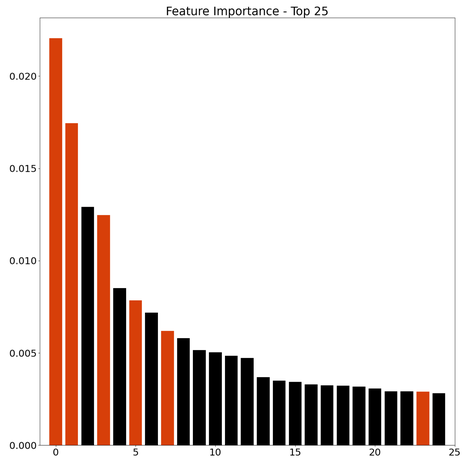

Figure 1: A representation of the top 25 selected features. Highlighted: features which represent valence/arousal.

We performed feature selection to determine the 25 features which best explained the variance in our data for the task of estimating valence and arousal. Three categories of features were found to best describe our data: (1) those which calculate a comment's valence/arousal, (2) those which determine valence/arousal of positive or negative words in a comment, and (3) those which calculate the intensity of anger.

Discussion

The previously discussed feature set was used to classify songs by mood. While some signal was present, performance leaves room for improvement, indicating that we may need to revise our feature generation model.

References

[1] Y.-A. Chen, Y.-H. Yang, J.-C. Wang, and H. Chen, "The AMG1608 Dataset for Music Emotion Recognition," in 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), South Brisbane, Queensland, Australia, Apr. 2015, pp. 693–697, doi: 10.1109/ICASSP.2015.7178058.

[2] S. M. Mohammad and P. D. Turney, "Crowdsourcing a Word-Emotion Association Lexicon," arXiv:1308.6297 [cs], Aug. 2013, Accessed: Nov. 04, 2020. [Online]. Available: arxiv.org/abs/1308.6297.

[3] Y. Choi and J. Wiebe, "+/-EffectWordNet: Sense-level Lexicon Acquisition for Opinion Inference," in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 2014, pp. 1181–1191, doi: 10.3115/v1/D14-1125.'

[4] A. B. Warriner, V. Kuperman, and M. Brysbaert, "Norms of Valence, Arousal, and Dominance for 13,915 English Lemmas," Behav Res, vol. 45, no. 4, pp. 1191–1207, Dec. 2013, doi: 10.3758/s13428-012-0314-x.