Complex Emotion Analysis of Song Lyrics

Darren Edmonds
Department of Computer Science, Johns Hopkins University

Abstract

Introduction

- Current sentiment analysis models fail to portray the coexistence of multiple emotions within a text sample, resulting in limited characterization of a human's true emotions[1]
- Songs are often created to elicit complex emotional responses from listeners, and can thus help us understand nuanced emotion[2]

This ongoing project examines a variety of approaches to the multi-emotion classification problem.

Model Selection

- Naive Bayes: widespread applications in sentiment analysis and text classification
- Random Forest: robustness to outliers; ability to deal with imbalanced data classification[3]

Dataset

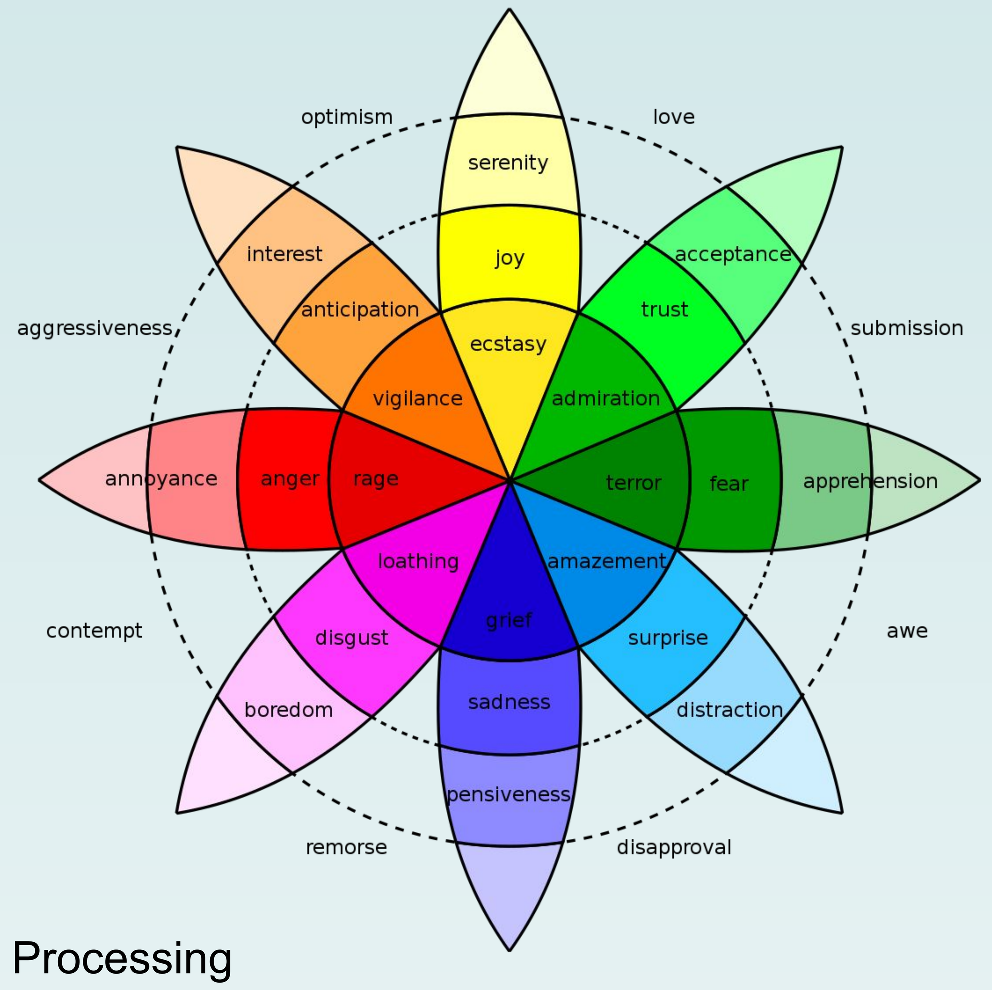
Consists of 600 song names, artists, and lyrics collected from Spotify, LyricFind, Genius, and MusixMatch.

Song	Artists	Lyrics	Joy	Trust	Fear	Surprise	Sadness	Disgust	Anger	Anticipation
Apollo	Hardwell, Amba Shep	So I can understand Fighting just to survive But you taught me I can We are the lucky ones We are, we are Oh we are the lucky ones We are, we are Just one day in the life	1	1	0	1	0	0	0	0
Lullaby	R3HAB, Mike William	Hypnotized, this love out of me Without your air I can't even breathe Lead my way out into the light Sing your lullaby Cherries in the ashtray Take me through the day I just gotta make you drunk in memory	0	0	1	0	1	0	0	0
Melody (Tip Of I	Mike Williams	You stare a little too long We do this dance every time You and I, you and I Yeah, we're up in the air Like the smoke from your lips But I can't find my words	1	1	0	0	0	0	0	1
Take Me Home	Cash Cash, Bebe Rexh	I'm falling to pieces But I need this Yeah, I need this You're my fault My weakness When did you turn so cold	0	0	0	1	1	1	0	0

Methodology

Basis for label selection: Plutchik's Theory of Emotion[4]

- All emotions are combinations of 4 pairs of bipolar emotions
- Labeling the presence or absence of these pairs can lead to additional emotion classification



Data Processing

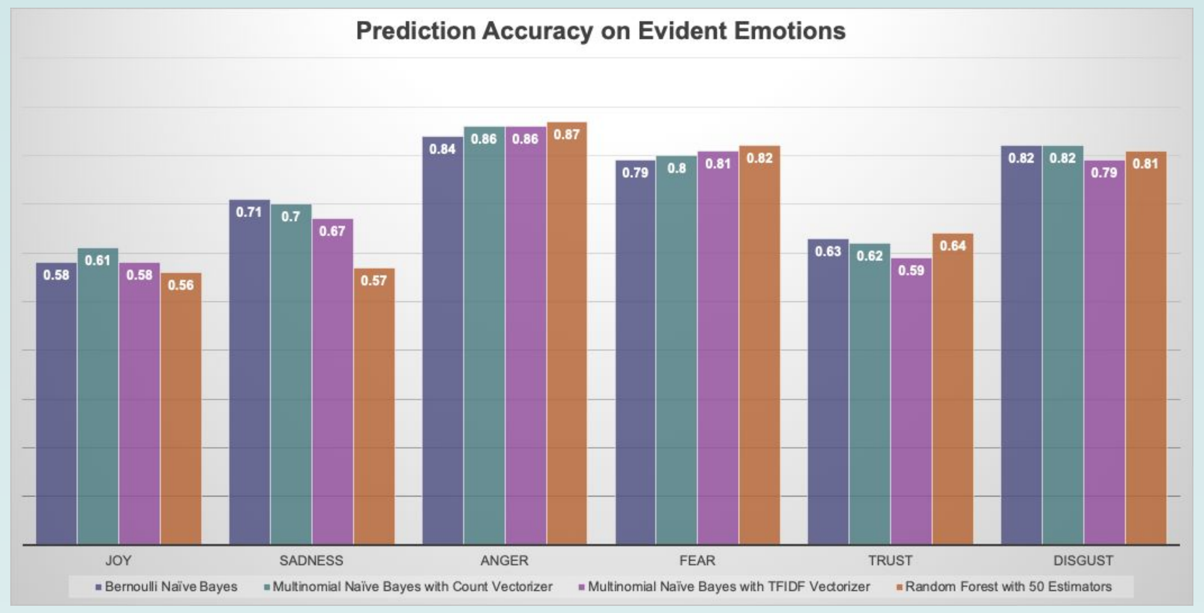
- Naive Bayes: Song lyrics were transformed into clean bag-of-words data through the removal of punctuation and capitalization
- Random Forest: Lyrical data used for Naive Bayes was transformed into a feature vector using the NRC Emotion Lexicon[5]

```
print(lexiconize("i will never abandon my wonderful abacus"))
[0, 0, 0, 1, 1, 1, 1, 2, 3]
```

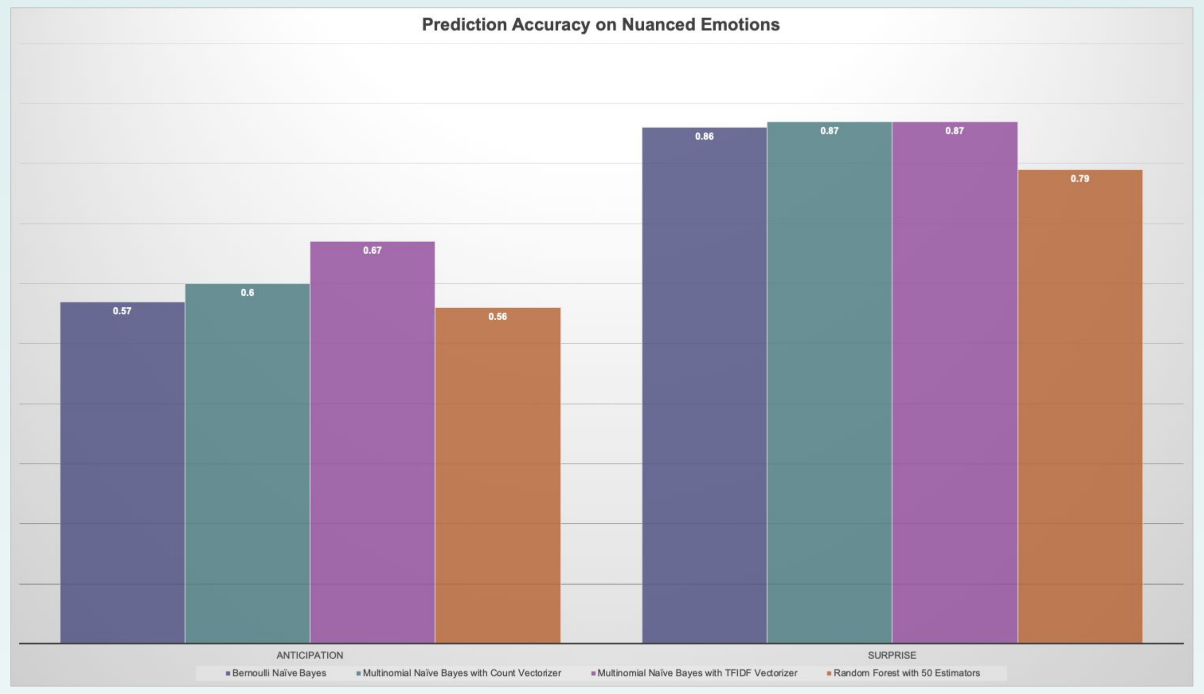
abacus	anger	0
abacus	anticipation	0
abacus	disgust	0
abacus	fear	0
abacus	joy	0
abacus	sadness	0
abacus	surprise	0
abacus	trust	1
abandon	anger	0
abandon	anticipation	0
abandon	disgust	0
abandon	fear	1
abandon	joy	0
abandon	sadness	1
abandon	surprise	0
abandon	trust	0
abandoned	anger	1
abandoned	anticipation	0
abandoned	disgust	0
abandoned	fear	1

Results

On average, emotions associated with negative sentiment were classified with greater accuracy than those associated with positive sentiment.



A multinomial Naive Bayes model with a TFIDF Vectorizer achieved the best prediction accuracy on anticipation and surprise.



Confusion matrix, F-measure, and Matthews correlation coefficient for selected emotions

Naïve Bayes (TFIDF): Anticipation	Absence Predicted	Presence Predicted
Absence Actual	39	19
Presence Actual	15	31

F1: 0.65, MCC: 0.34

Random Forest: Anticipation	Absence Predicted	Presence Predicted
Absence Actual	38	20
Presence Actual	26	20

F1: 0.47, MCC: 0.09

Naïve Bayes (TFIDF): Anger	Absence Predicted	Presence Predicted
Absence Actual	85	3
Presence Actual	12	4

F1: 0.35, MCC: 0.31

Random Forest: Anger	Absence Predicted	PresencePredicted
Absence Actual	87	1
Presence Actual	13	3

F1: 0.3, MCC: 0.33

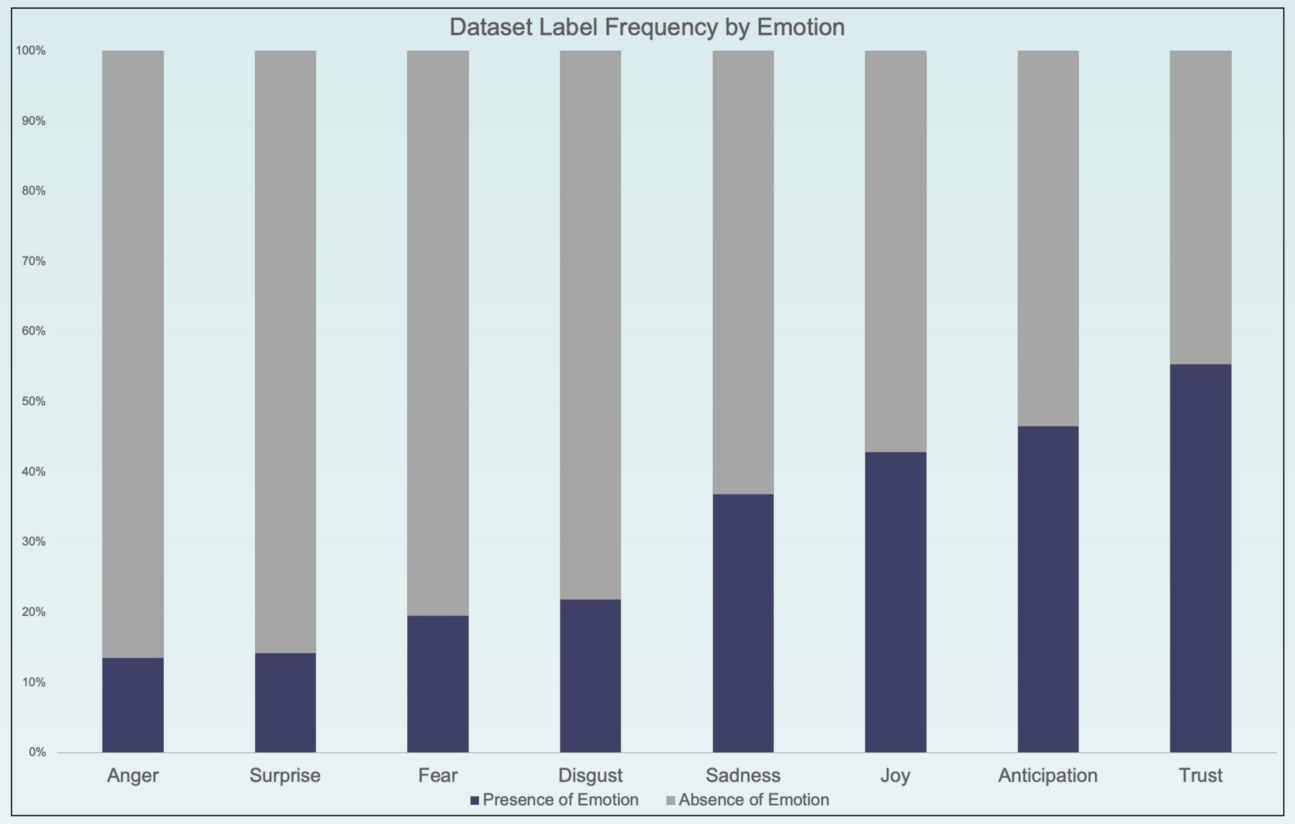
Conclusion

Optimal Models

- Naive Bayes (Multinomial, TFIDF): Classified surprise with >80% accuracy, sadness with >70% accuracy, and anticipation, joy, and trust with >60% accuracy.
- Random Forest (50 estimators): Classified anger, fear, and disgust with >80% accuracy.

Shortcomings

- Dataset labels based off of judgment of one person
- Tendency of both Naive Bayes and Random Forest models to disproportionately misclassify minority class



Future Work

Balancing Class Predictions

- Imposing higher-weighted penalty on misclassification of minority class
- Exploration of oversampling and undersampling techniques

Improving Dataset Quality and Size

- Active learning strategies
- Crowdsourcing of additional contributions

References

1. Robbert Jongeling, Proshanta Sarkar, Subhajit Datta, and Alexander Serebrenik. 2017. On negative results when using sentiment analysis tools for software engineering research. Empirical Software Engineering, 22(5):2543–2584.
2. Rada Mihalcea and Carlo Strapparava. 2012. Lyrics, music, and emotions. In Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, pages 590–599, Jeju Island, Korea. Association for Computational Linguistics.
3. Chao Chen, Andy Liaw, Leo Breiman, et al. 2004. Using random forest to learn imbalanced data. University of California, Berkeley, 110(1-12):24.
4. Robert Plutchik. 2001. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. American scientist, 89(4):344–350.
5. Saif M Mohammad and Peter D Turney. 2013. Crowdsourcing a word–emotion association lexicon. Computational Intelligence, 29(3):436–465.