

CMPU 366: Project Proposal

Adam Gulick

March 31, 2022

Proposal: Classifying Songs into Genre Based on Lyrics

For my Project, I wanted to do something similar to the [project](#) that was featured in one of the lectures earlier in the semester that analyzed the vocabulary of various rappers.

For my project I wanted to analyze and compare the lyrics of songs across various genres, rather than across various artists like the previous project. I would then like to attempt to predict the genre of a song when given a list of the words that are included in that songs lyrics. That is: Can you determine the genre of a song, given only a list of the words contained in its lyrics?

To do this, I will get the lyrics of a large collection of songs spanning a variety of genres in raw text form. Then, similar to the March 31 Lecture, spaCy can be used to tokenize and then some type of Classifier will be trained on these text files.

To test the prediction performance of the classifier, the lyrics of a new song can be given and the classier will attempt to put this new song into the correct genre based on the words contained in the lyrics.