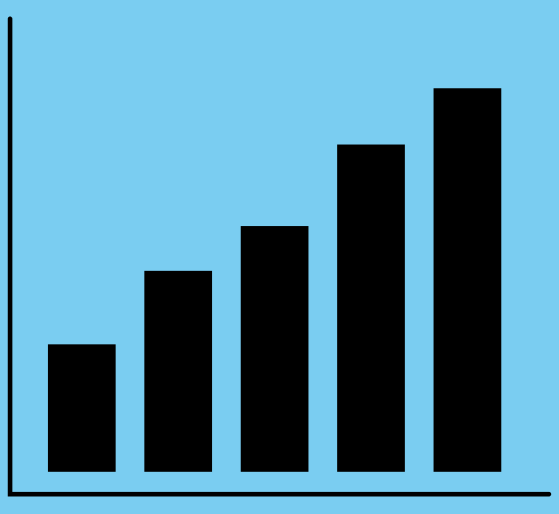
SENTIMENT ANALYSIS OF SONGS

Analyzing songs using ML

INTRODUCTION

- use characteristics of a song to create an analysis on the sentiment
- in similar projects, they solely focus on the lyrics, but we want to factor other portions of the song as well

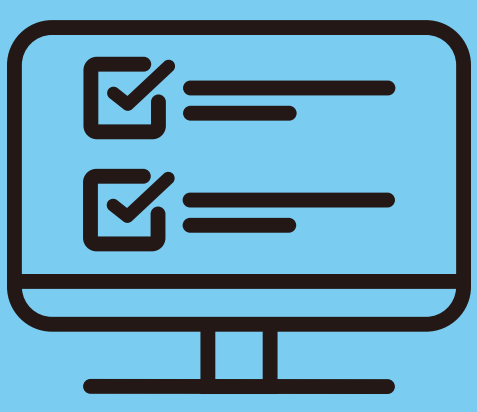


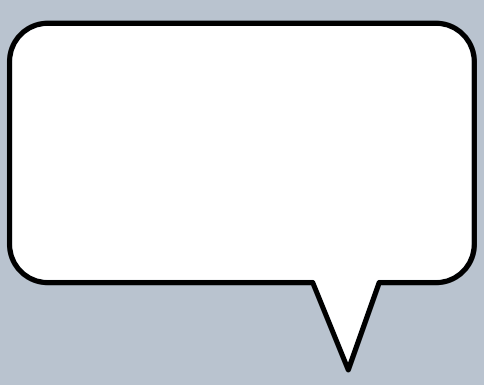
METHOLOGY

- Data set: Spotify and genius API
  - Data type: text data such as lyrics and audio data such as the song itself
  - Data set size: Unknown but will be the appropriate size
  - Targets: discrete
- approach
  - data preprocessing
    - Remove stop words
    - Apply text normalization
    - convert words into bigrams
  - supervised
    - use labeled data to train the model to predict the sentiment of the song
  - unsupervised
    - Use word2vec to convert bigrams into values
    - Use k-means with 3 clusters on the bigram vector

RESULTS

If successful then we will be able to gain a more accurate analysis of the song. To test our success, we will use accuracy to measure our success and use precision to check the quality of our results.





DISCUSSION

The risks involved is not setting up the data properly, which can result in misinformation. For our timeline, we plan on having the supervised model finished by the midterm report, and the unsupervised model done by the final report.