Classy Muse: Music Classification

Prachi Varma, Robert Turnage, Grant Wilson

Berkeley
UNIVERSITY OF CALIFORNIA

Problem and Objective

Problem Statement:

- Traditionally, music streaming services, such as Spotify, are provided musical genre metadata by labels and publishers. While this is generally reliable, it is not uncommon for metadata from smaller/indie publishers to be incorrect, leading to a bad user experience.

Objective:

- We want to build an ML model that will classify different musical genres by audio analysis
    - Defining the many community described music genres we will attempt to classify by analyzing audio features of frequency and time alongside metrics of track “energy”, tempo, and speed.

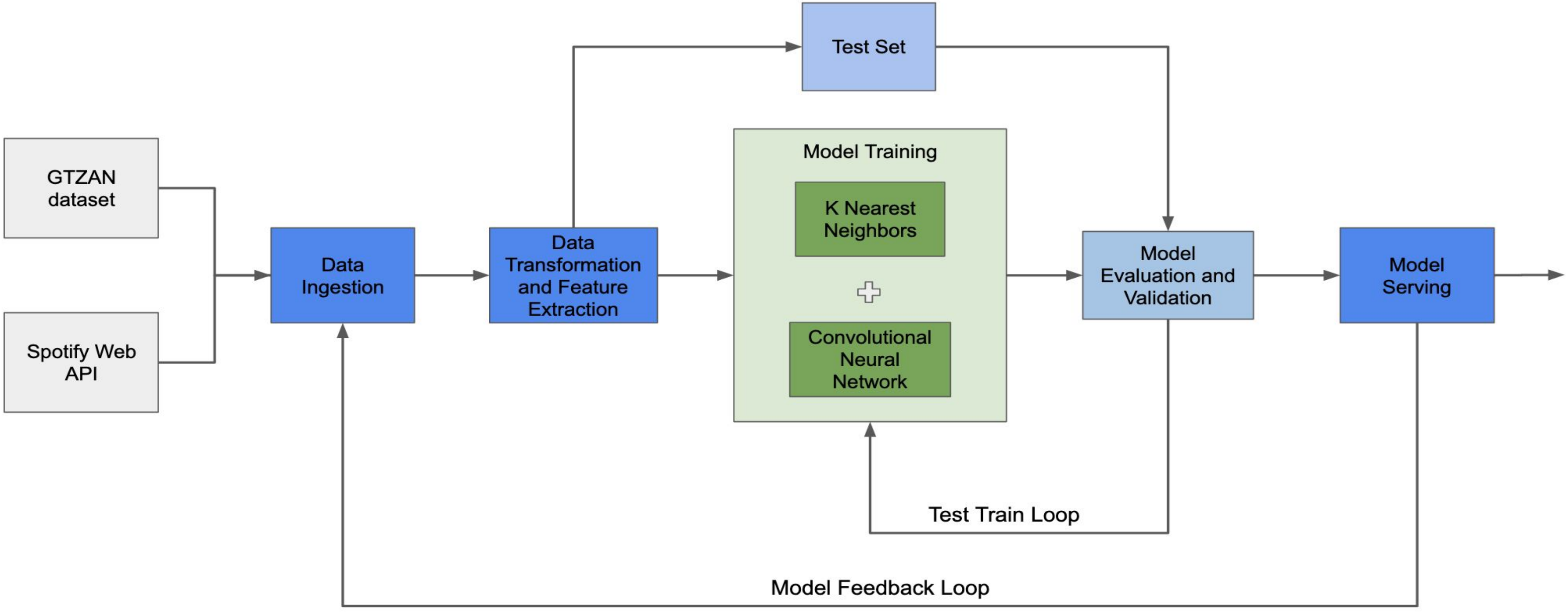
Berkeley
UNIVERSITY OF CALIFORNIA

Methodology

We want to train a model based on features that important in audio and specifically in music.

Audio Features

- Frequency
- Wavelength
- Amplitude
- Pitch patterns



Berkeley
UNIVERSITY OF CALIFORNIA

Data

1. GTZAN genre classification dataset (10 genres with 100 .wav files)
2. Spotify Web API/ Everynoise.com

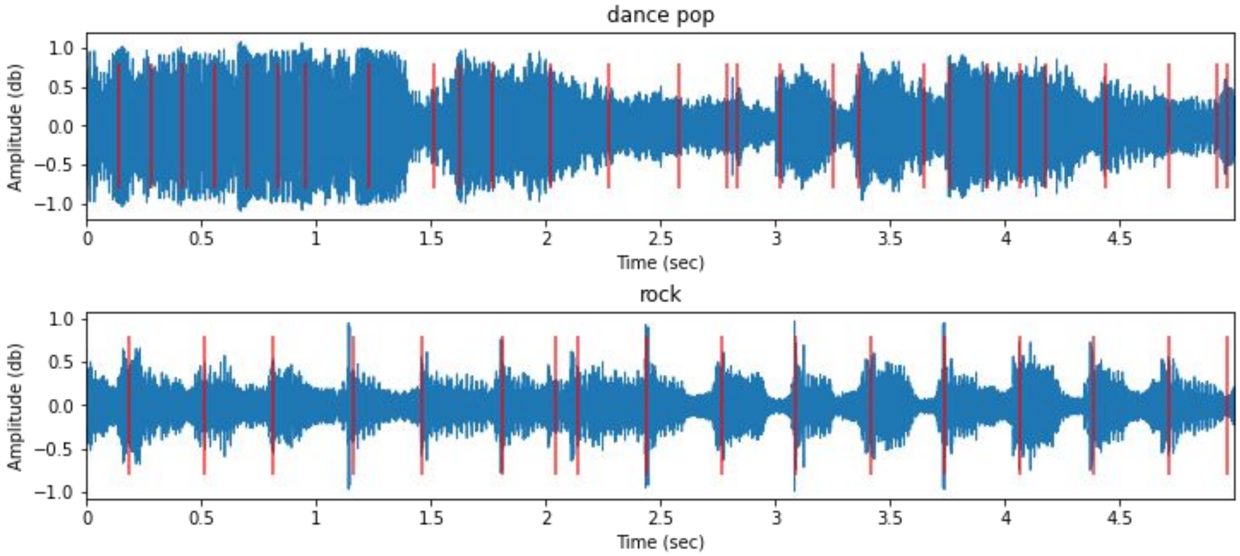
Berkeley
UNIVERSITY OF CALIFORNIA

EDA and Feature Extraction

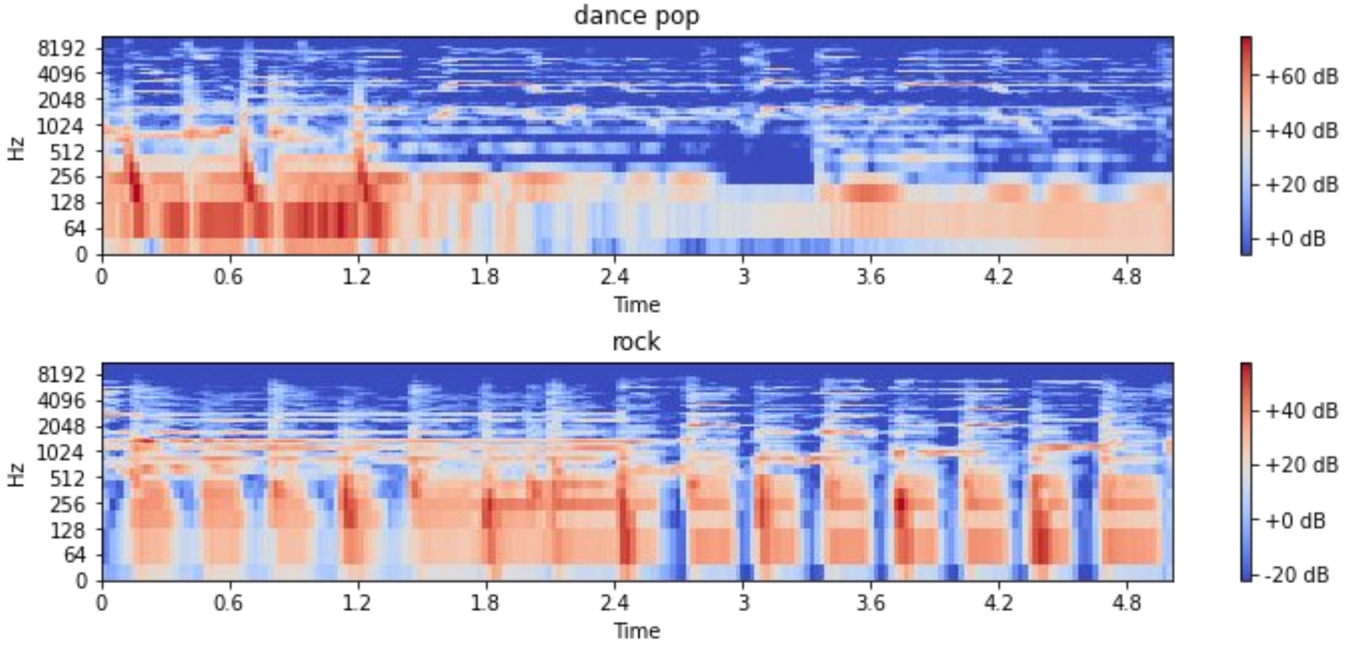
- Librosa for feature extraction from audio files
  - Waveforms
  - Mel-Frequency Cepstral Coefficients Spectrographs
  - Chroma Frequencies

Berkeley
UNIVERSITY OF CALIFORNIA

Waveforms

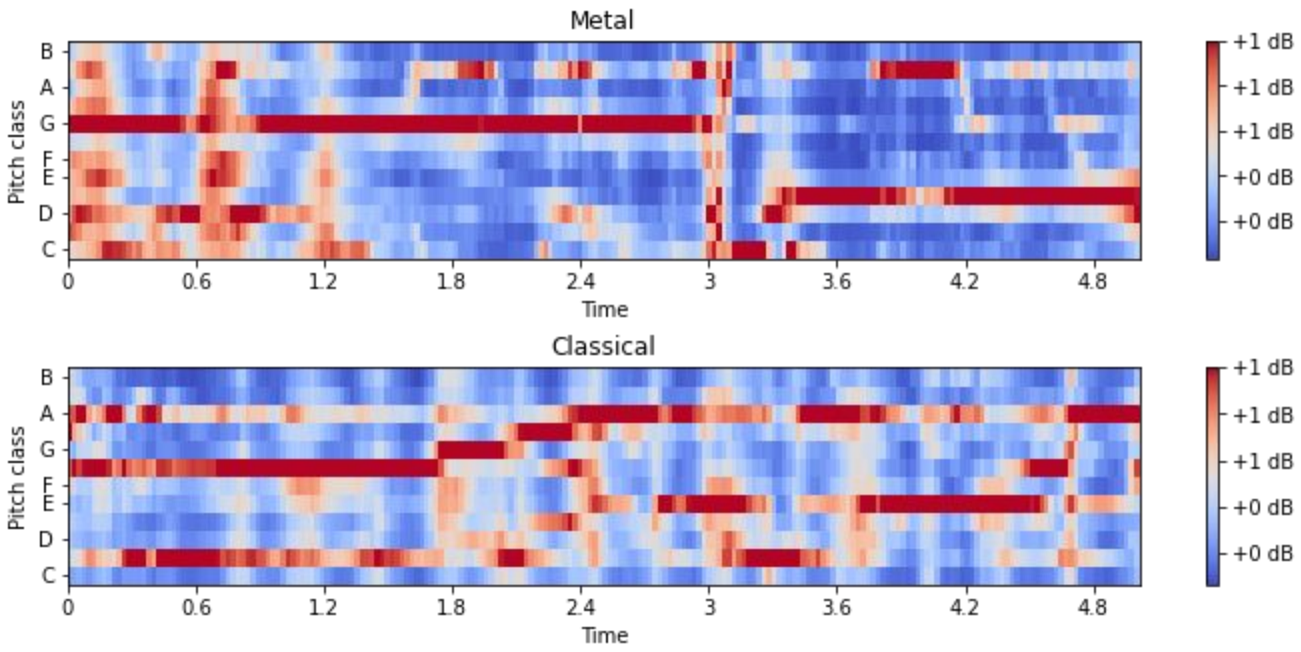

dance pop
Amplitude (db)
1.0
0.5
0.0
−0.5
−1.0
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
Time (sec)
rock
Amplitude (db)
1.0
0.5
0.0
−0.5
−1.0
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
Time (sec)

MFCC



Berkeley
UNIVERSITY OF CALIFORNIA

Chroma Frequencies



Berkeley
UNIVERSITY OF CALIFORNIA

Modeling

Convolutional Neural Network

- MFCC Spectrograph log transformation of data
  - Test Set: 400 files
  - Train Set: 100 files
- Constant Q Transform Chromagram
  - Test Set: 400 files
  - Train Set: 100 files

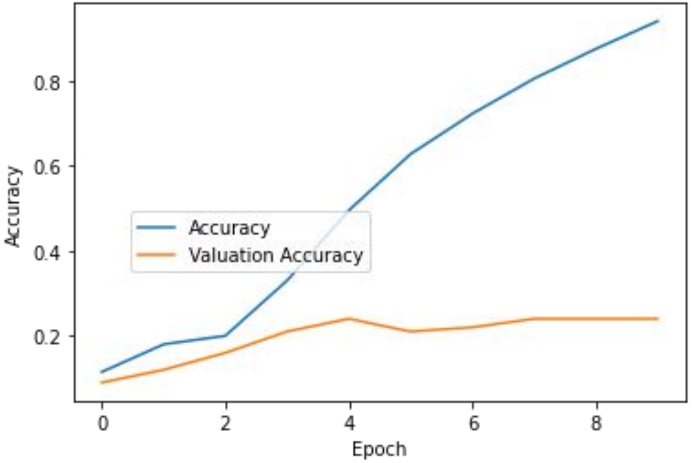
KNN, Logistic Regression, GPC

- DataFrame of 500 Spotify tracks

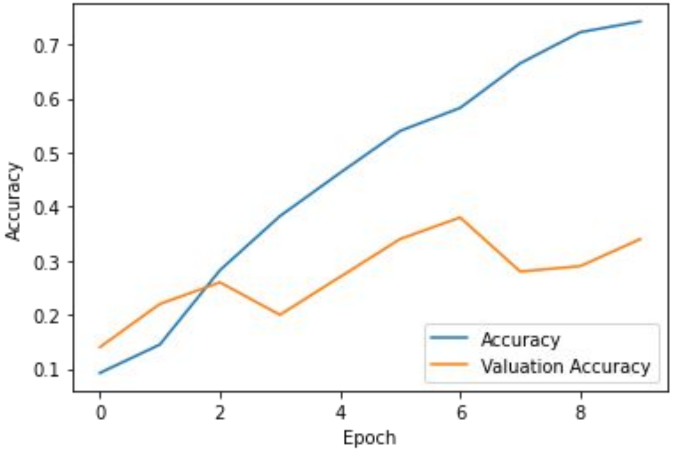
Berkeley
UNIVERSITY OF CALIFORNIA

CNN Results

CNN - Chroma CQT



CNN - MFCC Spectrogram



KNN Results

```
[ ]  search = GridSearchCV(pipe, param_grid, n_jobs=-1)
     search.fit(X_train, y_train)
     print("Best parameter (CV score=%0.3f):" % search.best_score_)
     print(search.best_params_)
```

```
Best parameter (CV score=0.307):
{'knn__leaf_size': 5, 'knn__n_neighbors': 10}
```

```
[ ]  print(f"Train Score: {search.score(X_train, y_train)}")
     print(f"Test Score: {search.score(X_test, y_test)}")
```

```
Train Score: 0.4272151898734177
Test Score: 0.29245283018867924
```

Berkeley
UNIVERSITY OF CALIFORNIA

Limitations and Future Work

- Limitations
  - 30 second audio snippets of much longer tracks may leave out significant features
  - Spotify doesn't allow queries based on genres, so data was pulled from a derivative of genre
    - Everynoise.com
  - Overfitting of the CNN due to the complexity of having multiple pooling and convolutional layers
- Future Work
  - Content-based music recommendation systems
  - Mood-based playlist creation

Berkeley
UNIVERSITY OF CALIFORNIA

Questions?

Appendix

Standards

- Librosa
- Scipy
- Sklearn
- tensorflow / Keras
- Pandas
- numpy