MUSIC-GENRE CLASSIFICATION

KELLY JARA

AUGUST 2023

Business Proposal

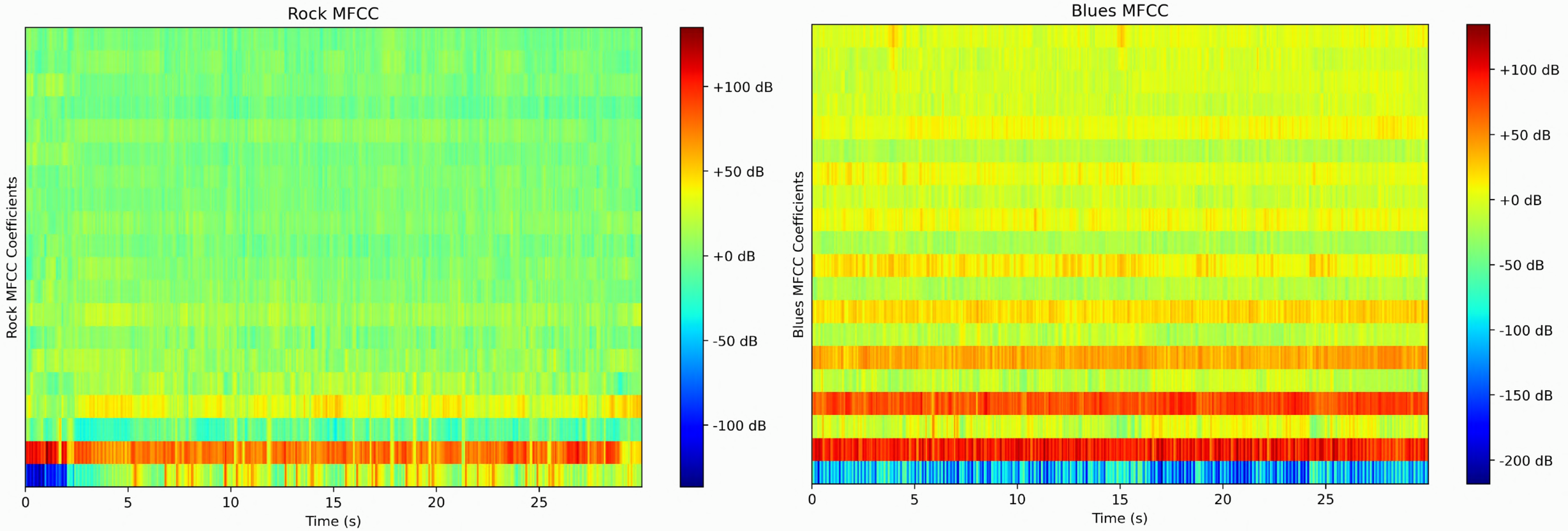
Streaming services have been around since the late 90's early 2000s. From services like Pandora to the now most popular Apple Music and Spotify, they all have something in common, playlist recommendations. At the base of those recommendations, how do these platforms know what to recommend to their users?

This is where music classification models come in.

How can we implement these models to improve recommendation systems?

MFCC

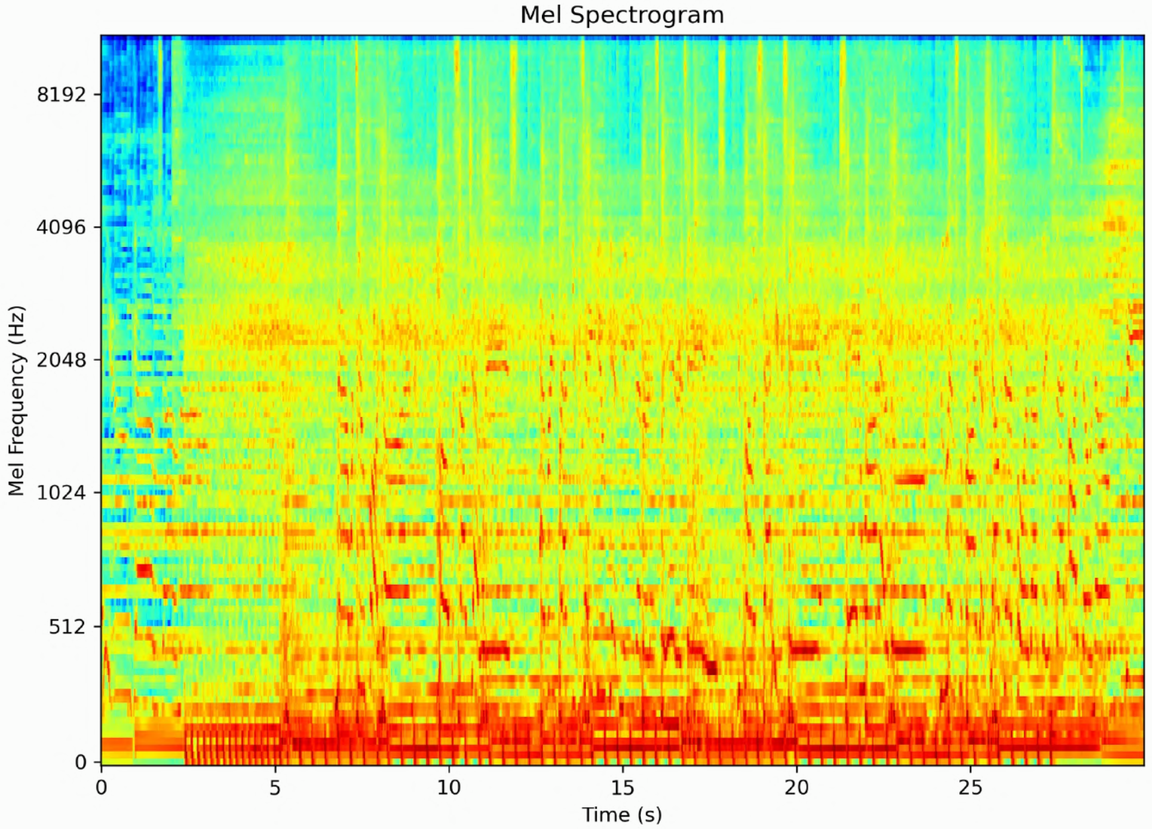
(MEL FREQUENCY CEPSTRAL COEFFECIENTS)

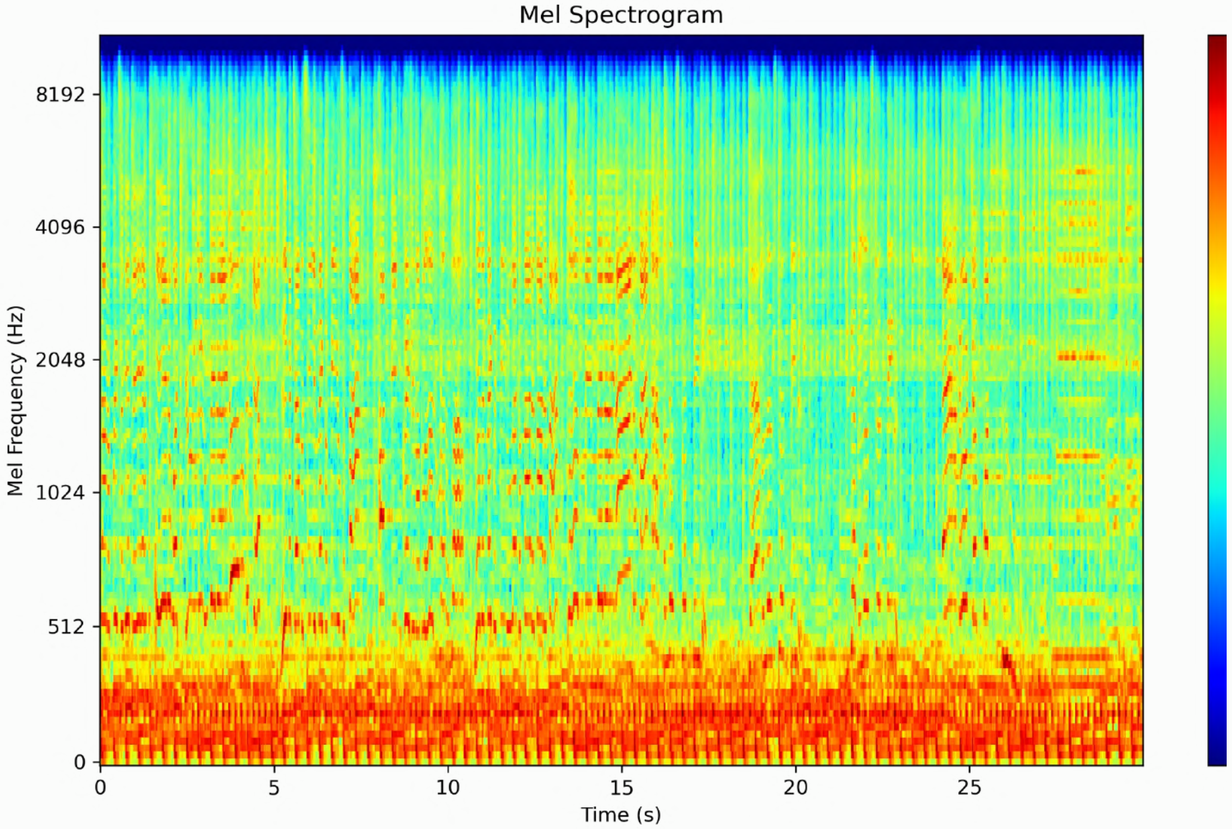


MFCC:

- Is made up coefficients are a representation of short-term power spectrum of sound.
- Is calculated with a discrete cosine transform(DCT).
- The bands within the spectrum are frequency bands that are spaced on the Mel-Scale.

MEL SPECTROGRAM

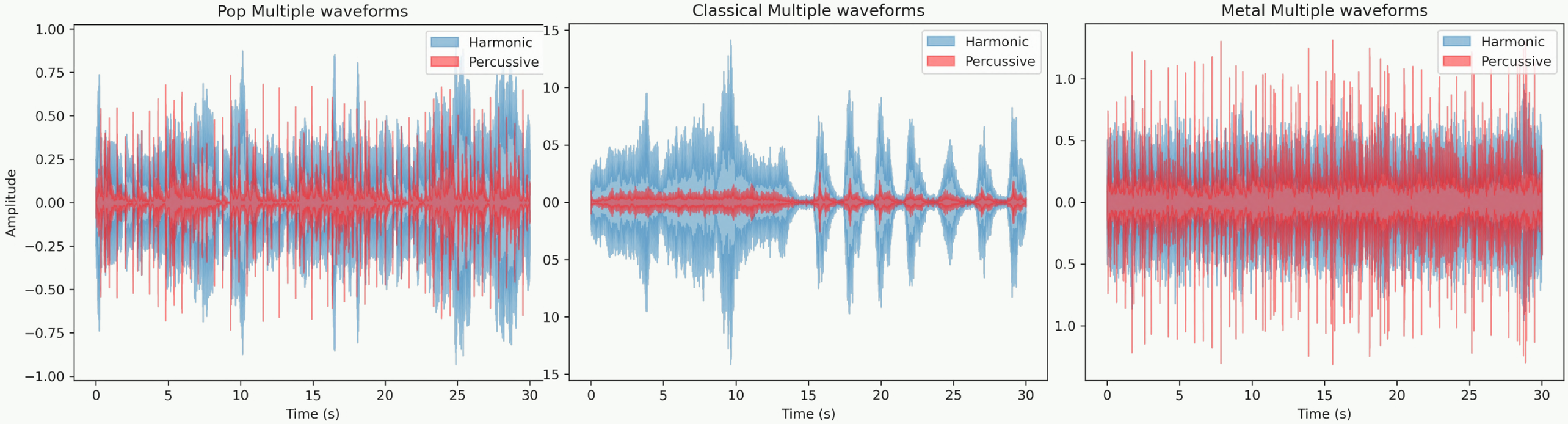




Melspectrogram:

- Uses a fourier transform in order to analyze the frequency content of a signal, which is then converted to Mel-Scale.
- Uses a linear space-frequency scaled.
- The bands, like mfcc, are based on the mel-scale.

HARMONIC & PERCUSSIVE WAVEFORMS



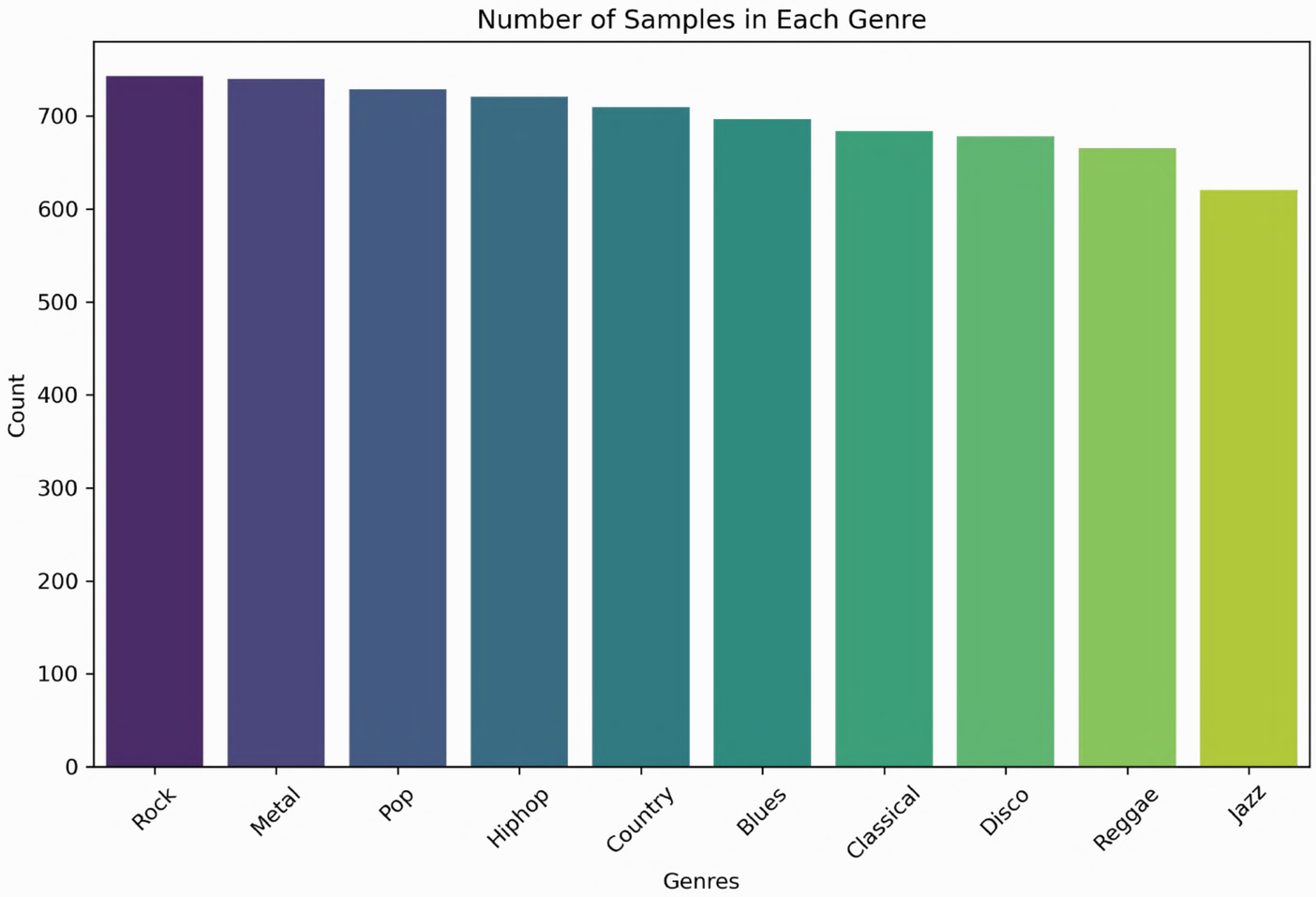
Harmonic Waveforms:

- are a type of waveform that consist of frequencies that are integer multiples of fundamental frequencies.

Percussive Waveforms:

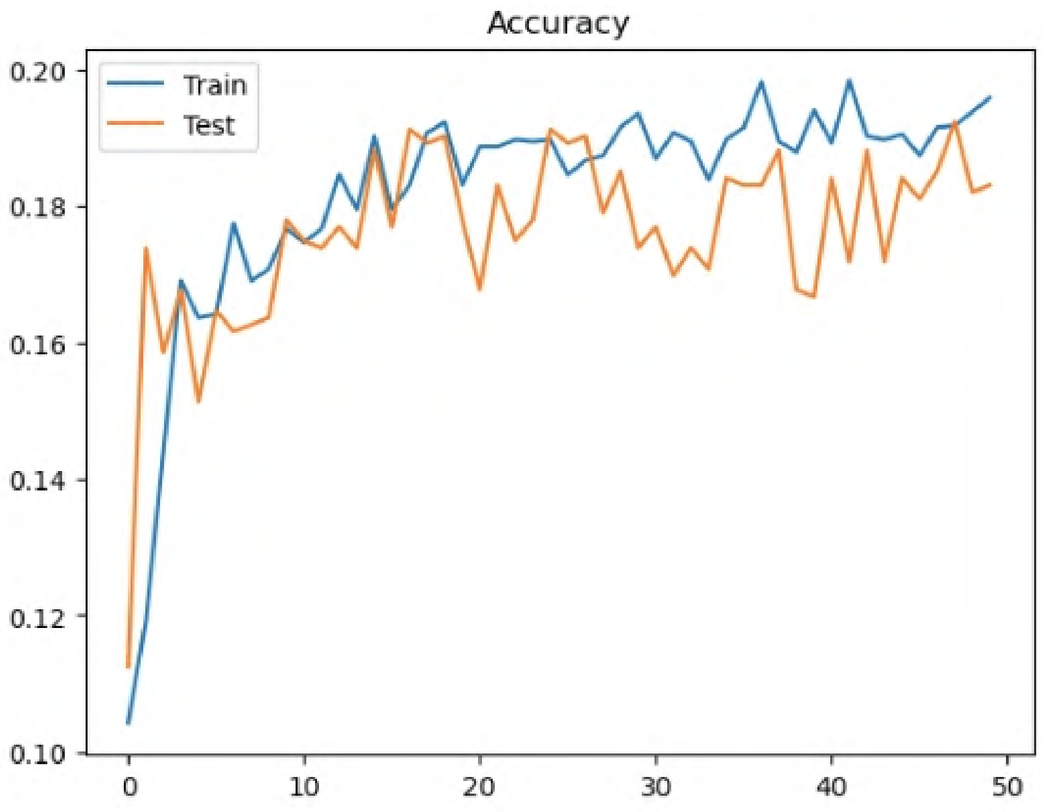
- consist of audio signals that are characterized by sudden, sharp and transient sounds.

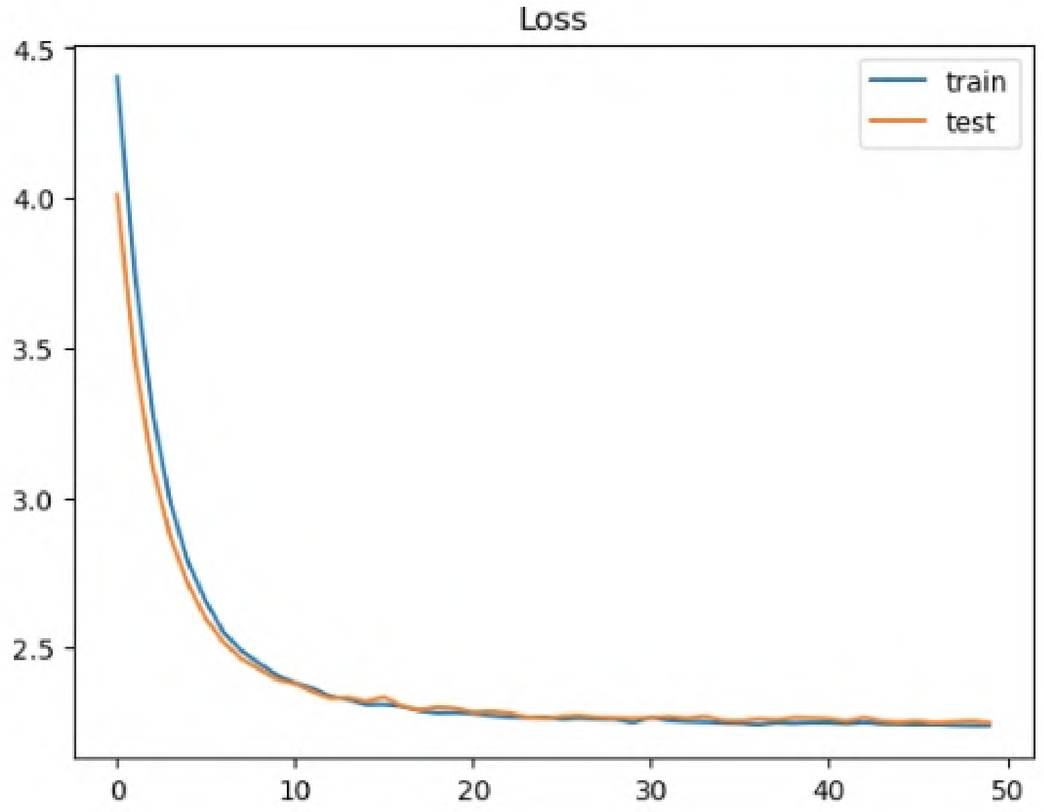
DATA OVERVIEW



- Audio samples from:
  - GTZAN Dataset (1,000 .wav files)
  - Personal audio files that were saved (800-1,000 .mp3 files)
- Models used:
  - Random Forest
  - Support Vector Machines (SVM)
  - K-Nearest Neighbor (KNN)
  - Recurrent Neural Networks (RNN) w/ Long-Short Term Memory (LSTM)

RNN WITH MFCC FEATURES





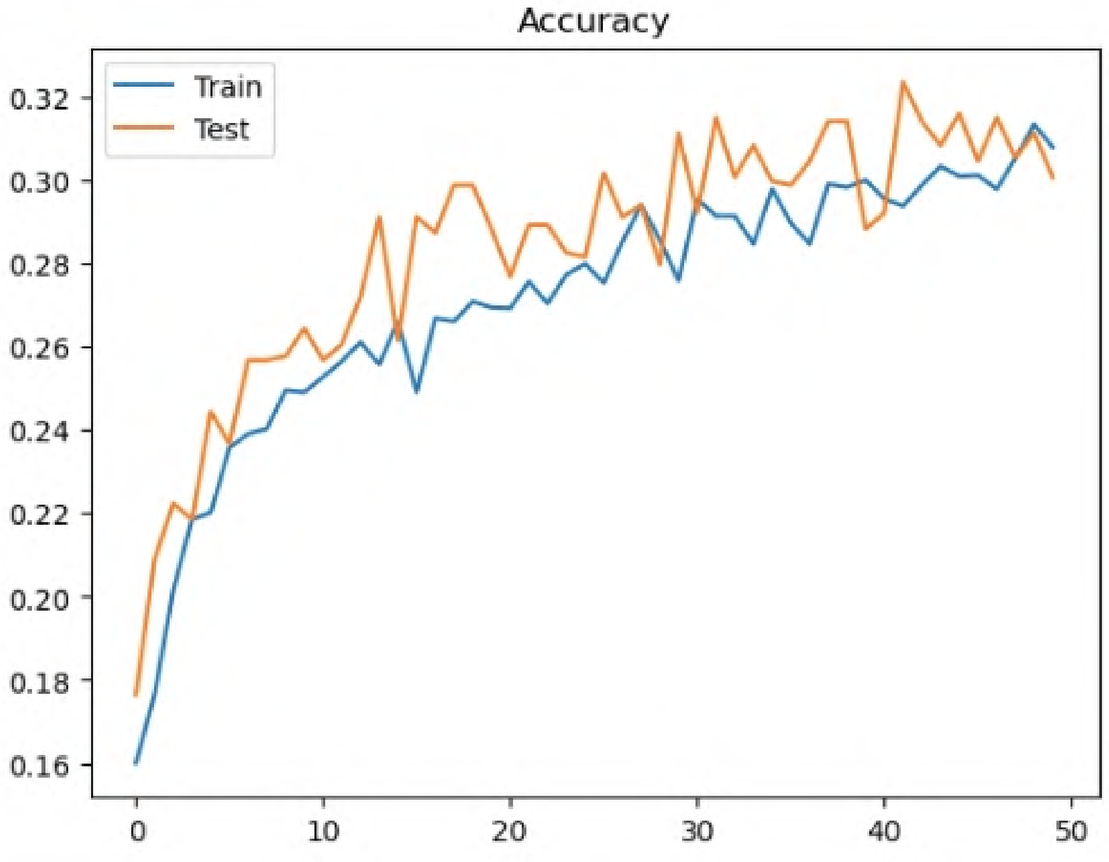
MFCC Features

- Sampling Rate (SR)
- Number of Mels
- Hop Length
- Fast Fourier Transform

Accuracy/Loss

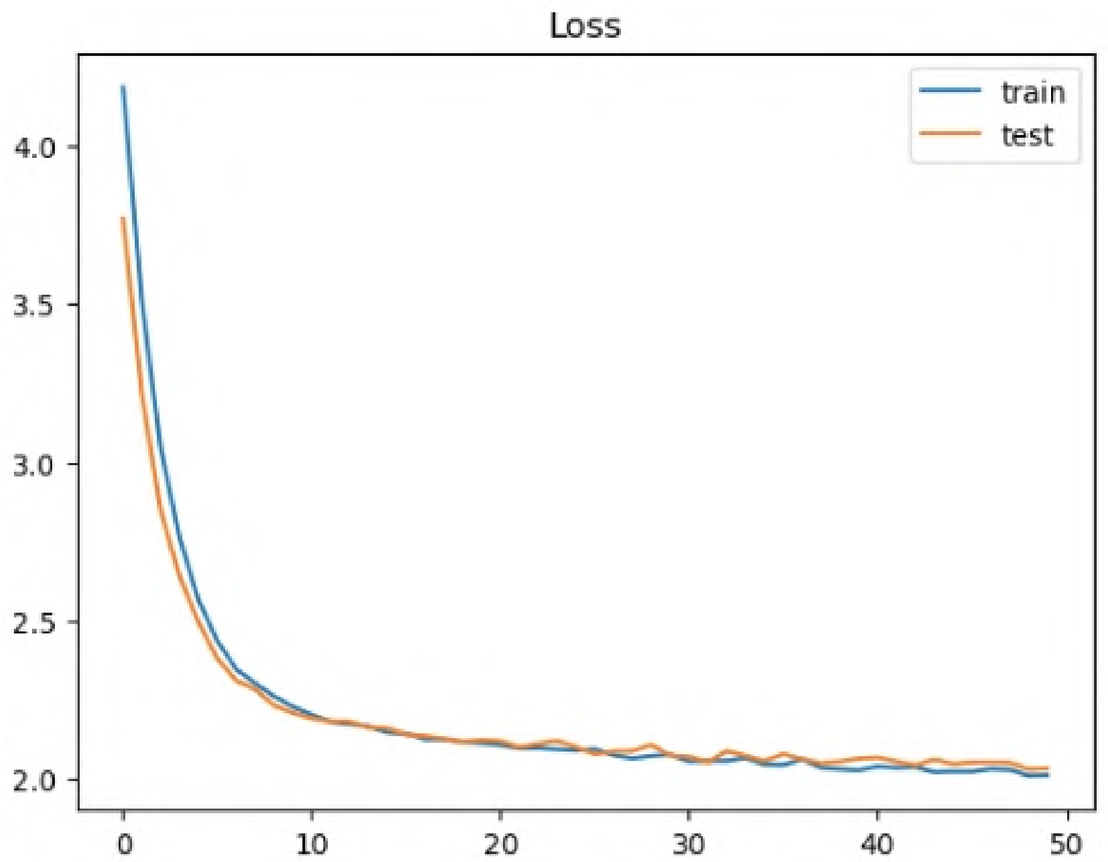
- Train:
  - Acc = 19.70% | Loss = 2.22%
- Test/Val:
  - Acc = 18.30% | Loss = 2.25%

RNN WITH MEL-SPECTROGRAM FEATURES



Mel-Spectrogram Features

- Sampling Rate (SR)
- Number of Mels
- Hop Length
- FFT



Accuracy/Loss

- Train:
  - Acc = 32.06% | Loss = 1.97%
- Test/Val:
  - Acc = 30.06% | Loss = 2.03%

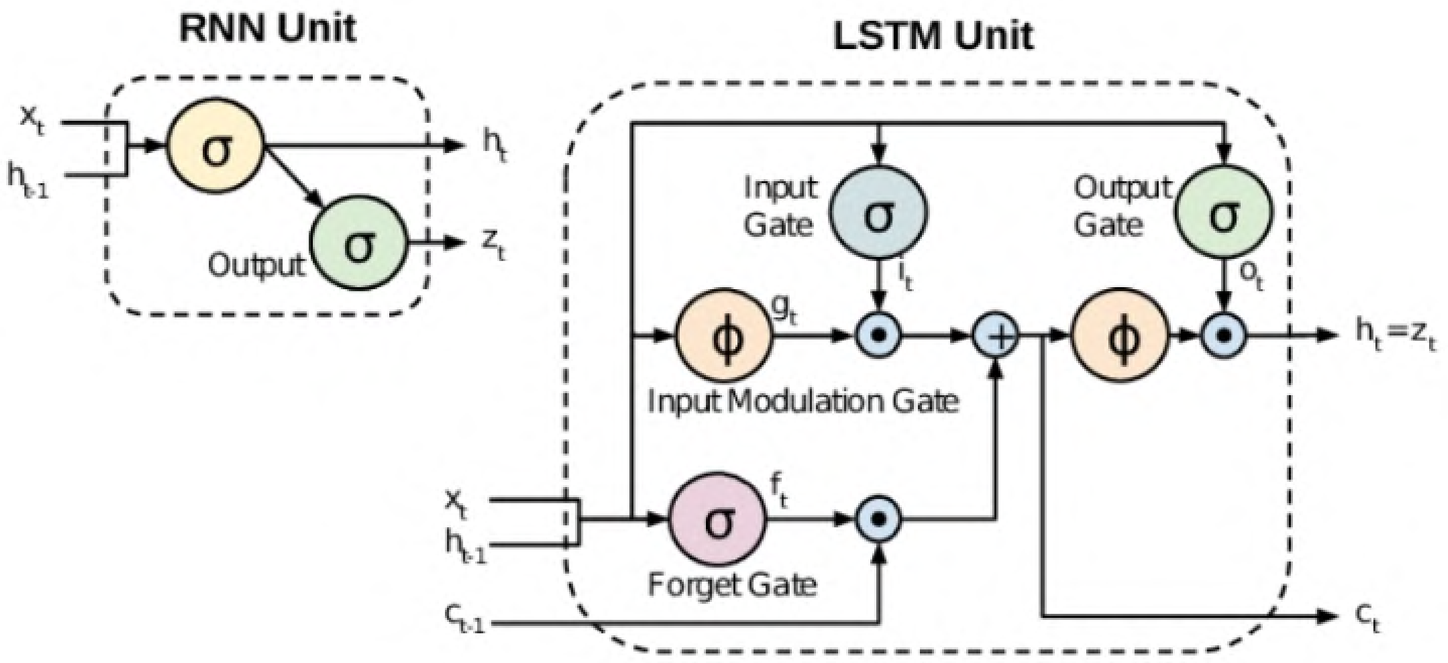
RNN MODEL

Recurrent Neural Networks are designed to process sequential data. RNNs have hidden layers that captures information from one output and uses that information in the next sequence of the model.

- Using an LSTM with the model allows it to hold information for longer periods of time. Long-short term memory.
  - Uses a gated mechanism that controls the flow of information in and out of the cells.
    - gates used: input gate, forget gate, and output gate.

ISSUES WITH RNN MODEL

- Lack of sufficient data
- Extraction of features
- # of features



KNN MODEL

Used 5-fold cross validation model that iterated over a set number of neighbors (1-99). The best parameters for the model were to have a number of neighbors equal to 1.

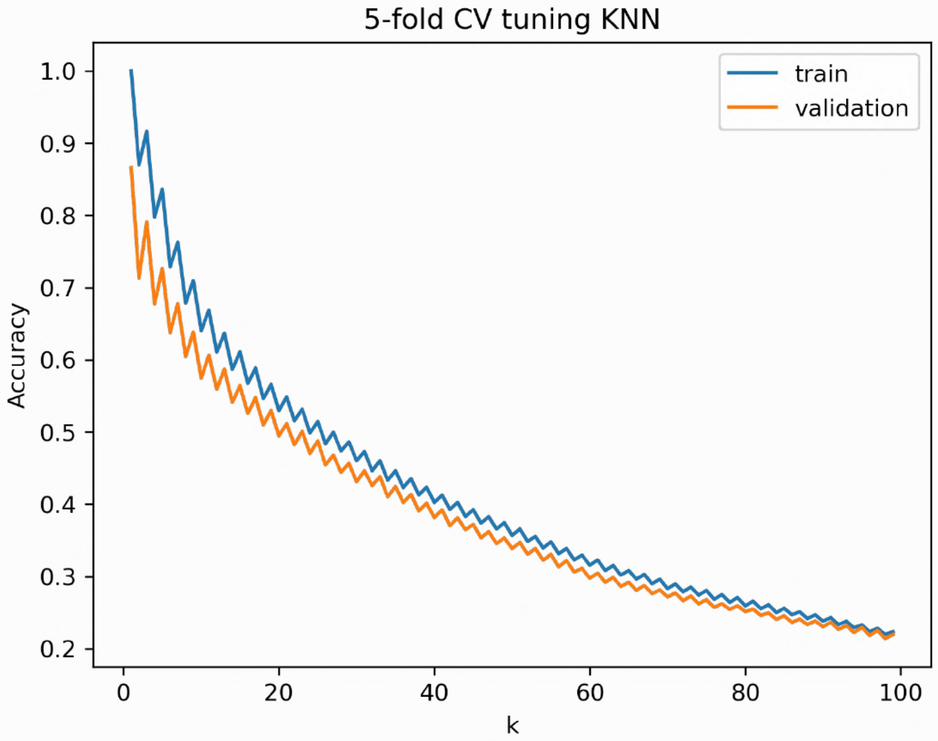
Precision = 89.24%

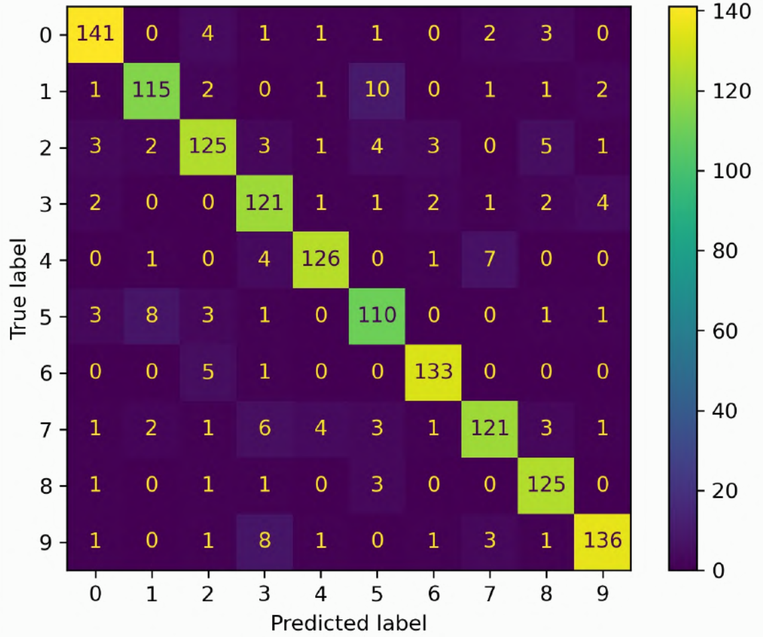
Recall = 89.10%

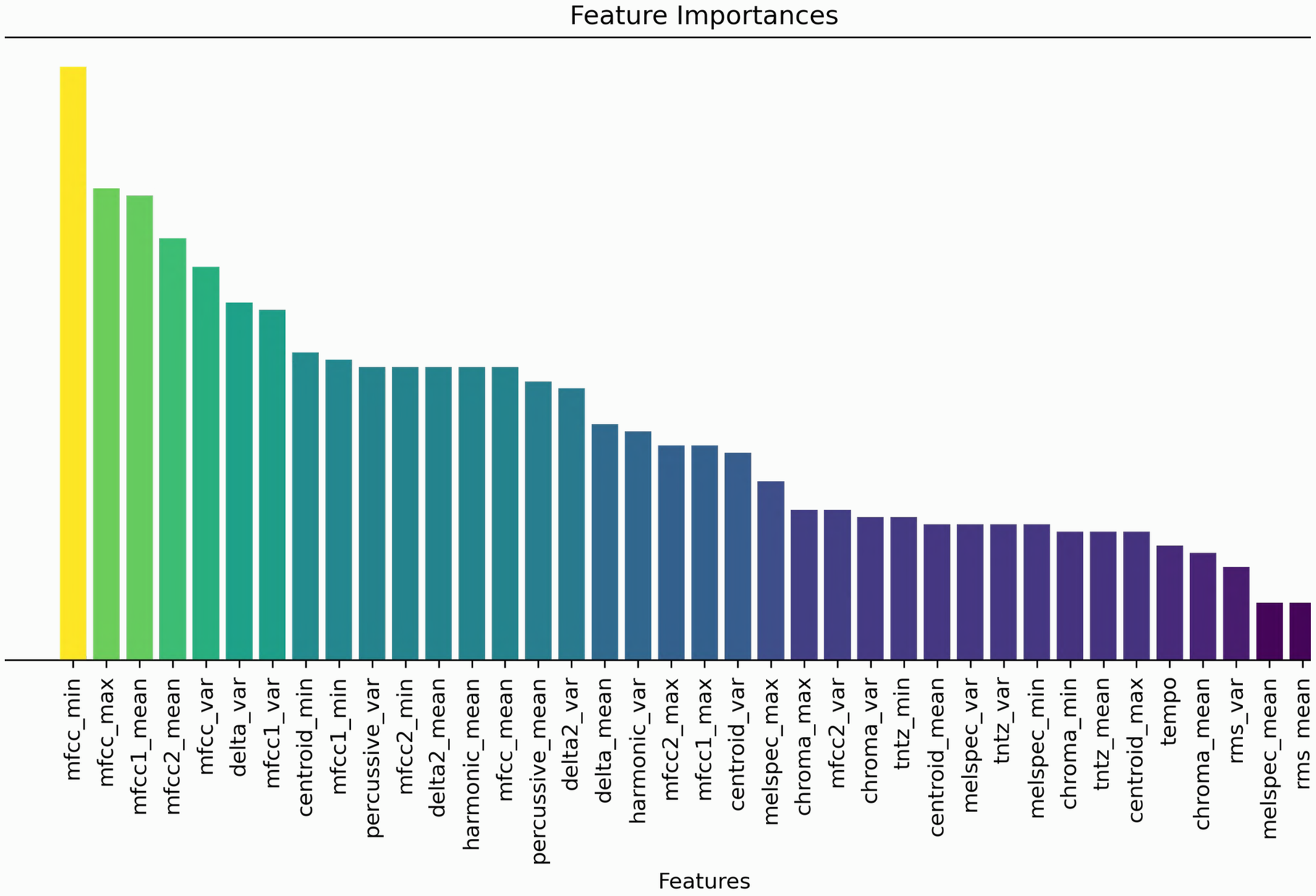
Accuracy = 89.13%

Genres Performing well:

- Blues
- Hiphop
- Metal
- Pop
- Rock





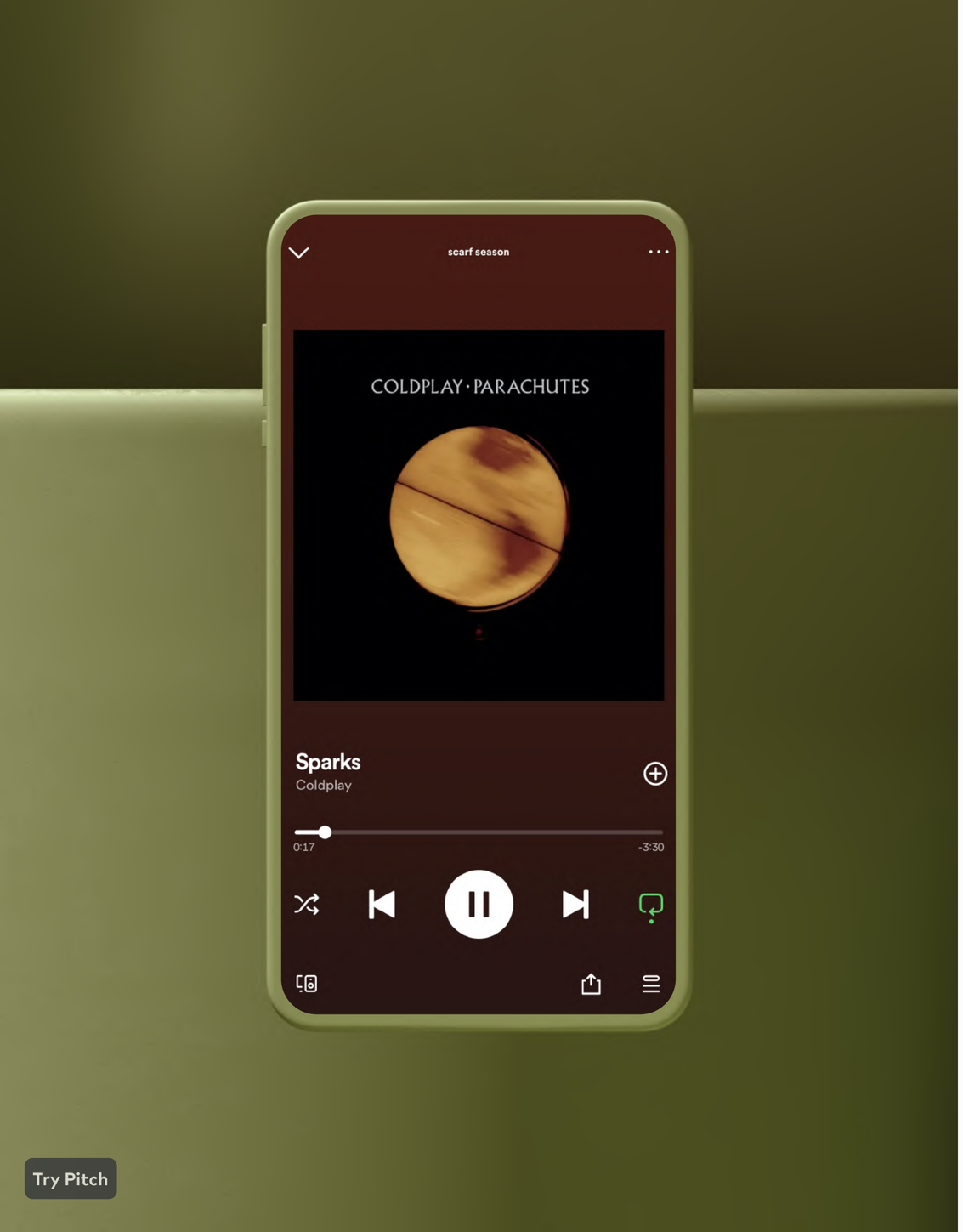


Features were extracted by getting the mean, standard deviation, min, and max values of the overall audio sample.

- Mfcc features take up the 10 top features
- Other important features:
  - Harmonic waveforms
  - Percussive waveform

IMPROVEMENTS

- Feature Extraction:
  - Experiment with different feature extractions such as:
    - Sampling rate
    - NFFT
    - Hop_Length
    - Frame length
- Improve RNN Model
  - Adding more layers
  - Obtain more audio samples
- Audio Samples:
  - Improve quality of audio files:
    - Limit audio from live performances
    - Audio file length (6 seconds vs 30 seconds)



Next Steps

- Deploy music classification app that allows users to input an audio file and receive an output of that audios genre to show accuracy of the model.

- Explore classifying more genres:
  - Explore music from different cultures and how it affects the performance of the model

- Create a genre-based recommendation system for user playlists.

Pitch

Want to make a presentation like this one?

Start with a fully customizable template, create a beautiful deck in minutes, then easily share it with anyone.

Create a presentation (It's free)