Capstone Project 2: Milestone Report

Springboard Data Science Career Track
Ishmail Grady
July 2019

Problem Statement

Music information retrieval is a growing field of study and with the advancement of robust machine learning techniques we can use models to classify music. Classifying music by genre can be used to provide content-based music recommendations or as the first step in determining what song is being played. Classifying music by genre is an important feature in any application that uses audio data or interacts with music files. Many different companies can benefit from a music genre classification models such as music streaming platforms like Tidal, Spotify, and Apple Music by allowing them to provide recommendations based on the music that they listen to.

The goal of this project is to create a neural network using Python that is designed to take audio features of a MP3 file and predict the musical genre of the track. This can be especially useful for streaming platforms that allow users to listen to their music files in their applications that may have incomplete metadata, such as Soundcloud. For applications where classify this exact track is the end goal, this model can be incorporated into a larger model/algorithm to increase its accuracy and efficiency.

Data Source

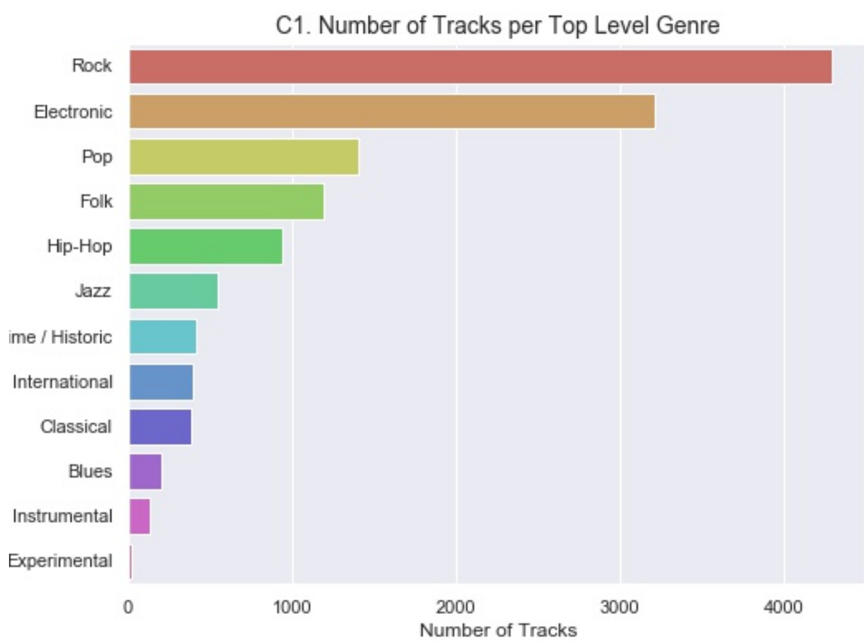
The main data source for this project will come from a compilation of high-quality audio files that are free and legal to use through the Free Music Archive which is designed for musical analysis. This dataset contains MP3 encoded 106,574 tracks from 16,341 artists and 14,854 albums that fall into 161 genres. The tracks are available as full-length files and 30 sec. snippets. The dataset contains metadata on each track and pre-computed audio features and is publicly available download. This dataset also offers audio features extracted from the Echonest API for a subset of 13,129 tracks. This subset will have 249 more audio features to choose from which could be an advantage. Considerations for dimensionality and predictive power will drive the decision for which version of the dataset to use.

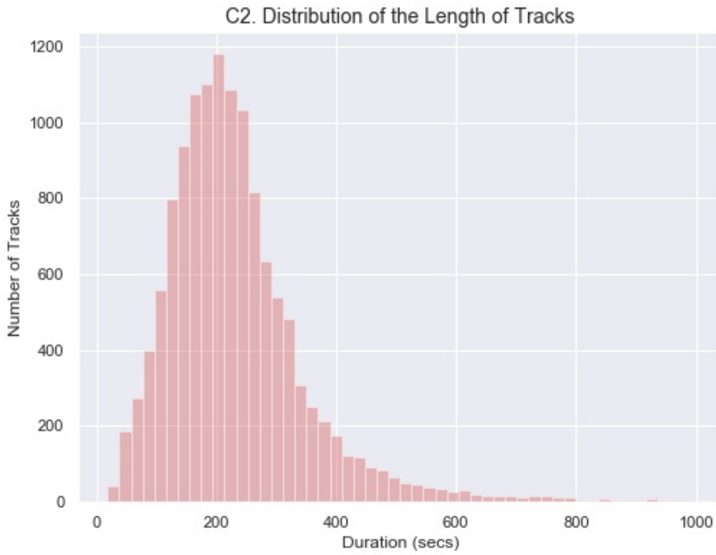
Data Wrangling & Cleaning

The raw data was fairly clean and tidy; each row represents a track and the columns include metadata for each trach and various audio features. The datasets for this project comes from 4 main tables: Tracks (track metadata and response labels), features (statistical audio features), genre (genre labels and hierarchy), and echnonest (audio features derived from Echonest API). Data contained in these tables was merged to produced two datasets: one containing Echonest features and one without. Some of the tracks in the dataset were not label with their top-level genre. For these, rows the top-level genre was inferred from another field that contained all of the applicable genres for the track (the highest in the hierarchy was chosen). The goal of this project will be predicting the top-level genre of a track based on its audio features, there for metadata such as artist name, number of listens, and year the track was made were removed from the dataset. Approximately 2200 tracks were missing labels and the top-level genre for these tracks was not able to be inferred and this data was discarded.

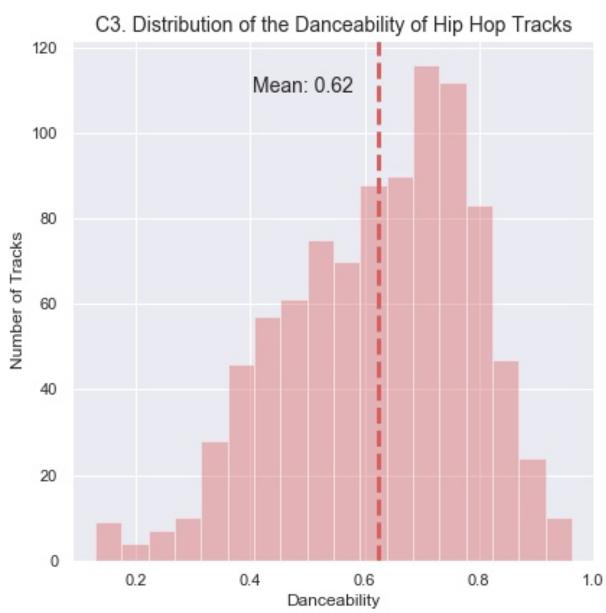
Exploratory Data Analysis

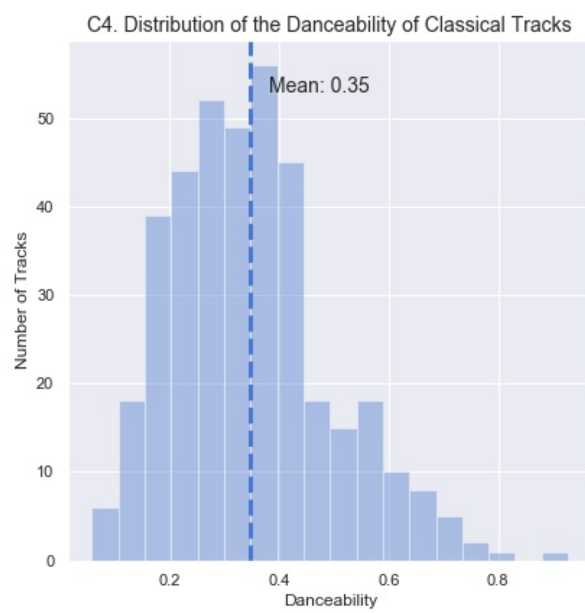
Understanding the proportion of tracks per genre in this dataset will be very important in developing a model. If the dataset is very unbalanced for certain genre's it will not perform well of a wide range of music. However, an equal proportion of track for each genre may not produce the best model given that we can expect some genres to be more popular than others. Figure C1 shows the number of tracks per genre for the echonest dataset. Rock and Electronic music have the largest proportions while Blues, Instrumental and Experimental represent a very small proportion of the dataset. These proportions indicate that that the dataset contains reasonable amounts of data from the most popular genres in the US according to a [2018 survey by Deezer](#).



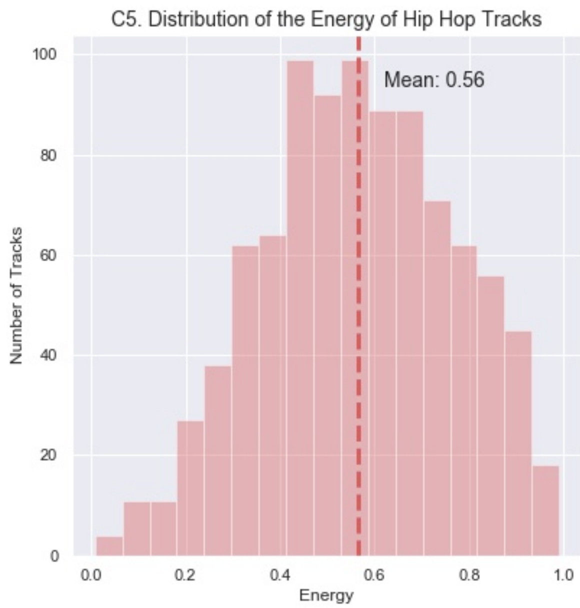


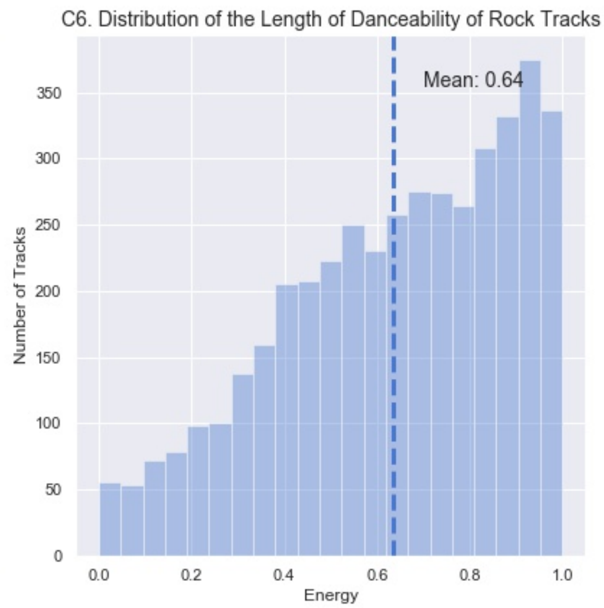
The Echonest API provides fields that quantify interesting qualities about music such as tempo, energy, and danceability. The use of these features should help the model different tracks from different genres. Figures C3 and C4 show the distribution of danceability ratings (a score from 0 to 1) for Hip Hop tracks and Classical tracks respectively. As expected, the Hip Hop tracks are rated higher in danceability than Classical tracks. In fact, a 95% Confidence Interval for Hip Hop track danceability yields a range of [0.61 - .63] while the interval for Classical tracks yields a range of [0.33 -0.36].



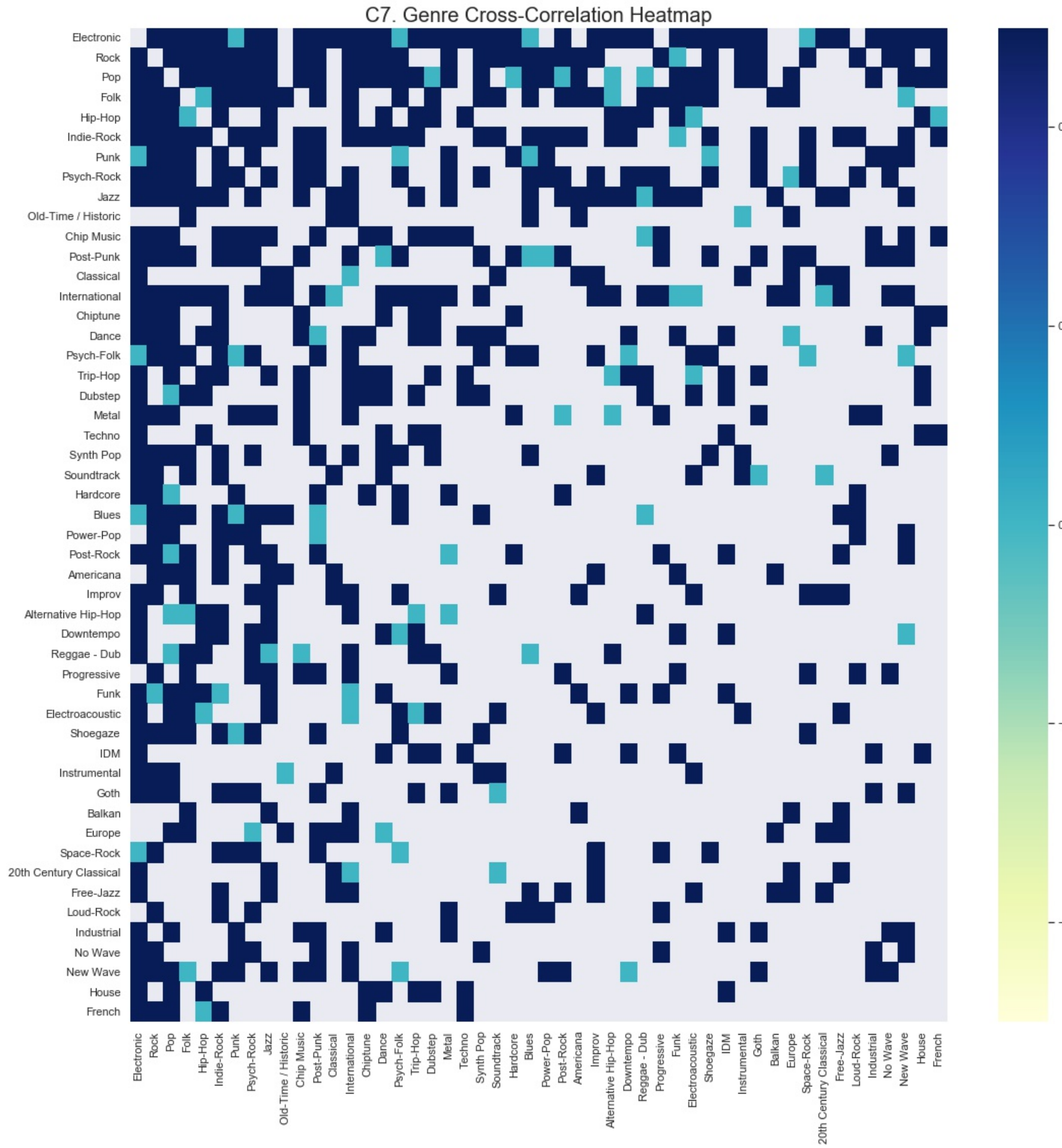


Similarly, the energy rating for tracks of different genres could increase predictive power. Figures C5 and C6 show the distribution of energy ratings for Hip Hop tracks and Rock tracks. Here, these distributions have quite different shapes. The distribution for Hip Hop tracks appears to be roughly symmetrical while the distribution for Rock tracks is skewed towards higher energy ratings. A 95% Confidence Interval for Hip Hop track energy yields a range of [0.55 - .68] while the interval for Rock tracks yields a range of [0.63 -0.64].





Discriminatory features will be very important given the amount of crossover between genres that can occur in the music industry. It is not uncommon for an artist to make music that can be classified as Hip-Hop, Pop, and R&B, artist Drake is a perfect example. Artist often collaborate across genres which can also mean the blending of sounds and creation of new sub-genres. Furthermore, genres such as Hip Hop can contain elements of many other top-level genres such as Jazz and Electronic. Figure C7 displays this cross correlation between genres and sub-genres.



C7. Genre Cross-Correlation Heatmap
Electronic
Rock
Pop
Folk
Hip-Hop
Indie-Rock
Punk
Psych-Rock
Jazz
Old-Time / Historic
Chip Music
Post-Punk
Classical
International
Chiptune
Dance
Psych-Folk
Trip-Hop
Dubstep
Metal
Techno
Synth Pop
Soundtrack
Hardcore
Blues
Power-Pop
Post-Rock
Americana
Improv
Alternative Hip-Hop
Downtempo
Reggae - Dub
Progressive
Funk
Electroacoustic
Shoegaze
IDM
Instrumental
Goth
Balkan
Europe
Space-Rock
20th Century Classical
Free-Jazz
Loud-Rock
Industrial
No Wave
New Wave
House
French
0.08
0.04
0.00
−0.04
−0.08