ORIE 4741 Project Midterm Report

An Insight into Sounds from Spotify

Vidita Gawade (vag39), Shruti Sanghavi (sjs449), Feiyu Tao(ft99)

October 27, 2017

I Introduction

Spotify is a music, podcast, and video-streaming services that was founded in 2008. In this report, we aim to build a tool that can predict the likeability of a song for an individual Spotify user. While Spotify does currently recommend songs to its users based on previous songs he or she listened to, the algorithms it uses focus on grouping similar music listeners together, also known as collaborative filtering. Our tool, instead, aims to use the features of songs among different genres to help determine an individual's likeability of a song.

II Data Collection, Creation, and Connection

As a first step, songs selected by the team were manually grouped into four genres. Each genre was then separated into two groups (Like or Not Like). In total, our database has eight playlists where the genres are Pop/Dance/EDM, Rock, Mellow and Hip-Hop/Rap. Within each playlist, there were approximately one hundred songs, ensuring an equal representation of songs per genre, per likeability.

Using Spotity's Developers toolkit, we created a user authentication token to approve access to the application. The API was connected via Python using the library: **spotipy**. Within Python, the eight playlists were combined, and the features "likeability" and "genre" were manually added since they were not part of the original Spotify features.

III Data Features

After cleaning the data, our dataset contains $n = 710$ data points and $d = 13$ features, which we use as our **initial** run. The audio features (the ones other than "likeability" and "genre") for each track were extracted using the API. The list of the features includes: *Acousticness*, *Danceability*, *Energy*, *Duration*, *Instrumentalness*, *Key*, *Liveness*, *Loudness*, *Mode*, *Speechiness*, *Tempo*, *Time Signature*, *Valence*, *Genre*, and *Like or not like*. In this report, only three features will be shown in details as an example. For a further description of all the features, refer to Spotify's audio feature description: `https://developer.spotify.com/web-api/get-audio-features/`. For the three selected features, their interpretation is included in the following table.

Features	**Interpretation**
Valence	Measures the degree of positivity of the song
Genre	Our dataset spans four basic genres mentioned above
Like or not like (binary)	1: the user likes the song; -1: the user does not like the song

IV Model Analysis

Since we want to categorize our predications into songs that "I like" or "I do not like", we decided to first try running the perceptron algorithm to see if our data linearly separable. As expected, the algorithm did not converge even if we allowed the misclassification error as high as 50%. The dataset seems not to be linearly separable.

Next, we tried to use the linear regression model to see how we could predict two things: the valence and the likeability of a song. For this, features were picked based on our logical intuition. For example, time duration was left out because it is not going to affect the valence of a song. The test-set misclassification error for predicting the valence of a song was 35.9%, and for predicting the likeability was 47.18%. These errors are higher than expected, based on a simple model used at this point, but they give scope for improvement once more advanced algorithms are used.

V Data Visualization: Insight into Spread of Features within a Genre

We were interested in seeing the spread of features within each genre. We focused on valence, acousticness, and energy.

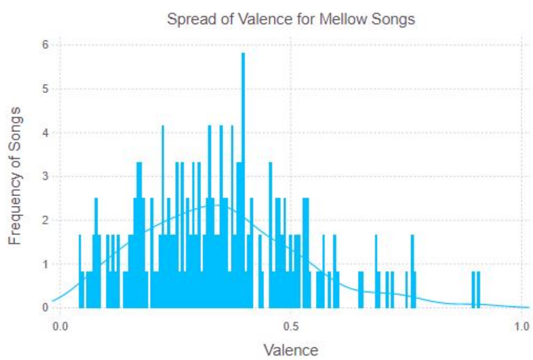

Figure 1

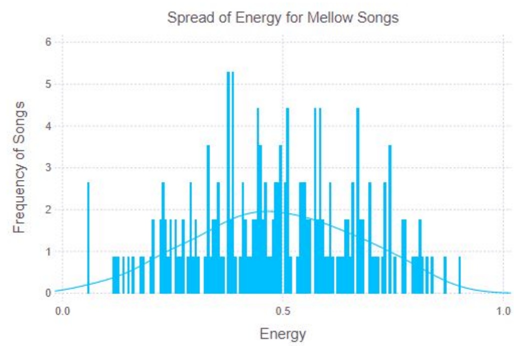

Figure 2

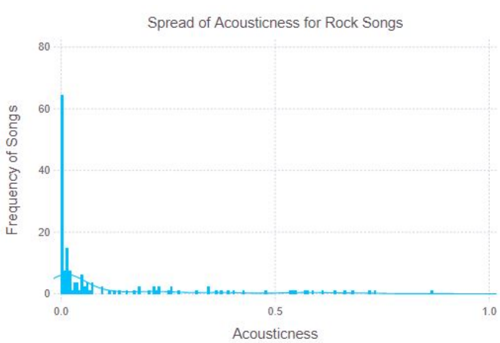

Figure 3

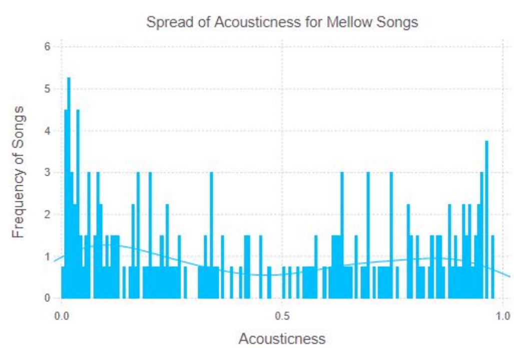

Figure 4

Energy describes how upbeat a song is (the closer a song is to 1, the more energetic it is). As expected, histograms (not shown) verified Pop/EDM/Dance, Hip-Hop/Rap, and Rock to be more energetic. The histograms (Figure 1 and Figure 2) for Mellow songs showed surprising results; energy levels were spread out, implying what an individual considers as a mellow song can in fact have happier and upbeat sounds associated with it.

Acoustic sounds use instruments to produce music as opposed to electronic means. As expected, histograms (not shown) verified Pop/EDM/Dance and Hip-Hop/Rap songs, songs created in studios, were much more electronic. But surprisingly, the histogram (Figure 3) for Rock songs, usually associated with acoustic sounds, showed stronger association with electronic sounds. Similarly, the histogram (Figure 4) for Mellow songs, usually associated with acoustic sounds, showed a rather wide spread range acoustically, favoring neither acoustic sounds nor electronic sounds.

VI Future Work and Consideration

VI.A Future Feature Transformation

Currently, each value of our genre feature corresponds to an integer value from 1 to 4. This makes our data more categorical than numerical, and incorrectly implies that one genre holds more importance over another genre. Changing the numerical value of the labels should actually give us the same predictions, but for models such as linear regression, this is not the case. In order to deal with this situation, we have decided to utilize feature transformation and convert the "genre" feature into dummy variables. We would create a vector for each "genre" value. For example, (1,0,0,0) might refer to genre 1, and (0,1,0,0) might refer to genre 2. In this way, the final prediction would be independent of the notation of the genres.

VI.B Avoiding Overfitting & Underfitting

A couple of strategies that we will be using to avoid over/under-fitting situations are as following:

1. We will collect more songs so that there will be more data available to us.
2. Although we used logical intuition to deselect features in our preliminary analysis, we plan to use advanced feature selection to avoid building a complex model to not overfit.
3. While we have set up the training and testing infrastructure, we would like to build cross validation to ensure that the models we compare and eventually select are robust to use in the wild.

VI.C Future Work

The results show potential for building a tool that can predict the likeability of our song. In addition, since the data was directly collected from Spotify, it is not missing or corrupted. We want to explore if it is better to predict the likeability of songs as a binary classification, or as a percentage of likeability. We plan to apply binary classification models such as logistic regression, decision trees, random forests, bagging, and support vector machines. If we plan to use percentage of likeability, we plan to explore linear regression and other feature transformations for our categorical features, in addition to performing regularization. Through this, we aim to use the rich features of a song and abilities of different models to help to predict the likability of a song for an individual.