Music Genre Classification: Predicting Neo-soul Subgenre Using Logistic Regression

Chris Hyorok Lee
School of Jazz and Contemporary Music, The New School, New York, USA
hyorok.lee@gmail.com

Abstract

Making into a Spotify Official playlist is one of the key steps for a successful album or a song release for any songwriter or producer in the music industry. This project investigates a way to use machine learning techniques and data from Spotify API to predict whether a certain song will fit a target Neo-soul Spotify Official playlist, a specific subgenre of RnB, specifically using Logistic Regression, K Nearest Neighbor (KNN), and Support Vector Machine(SVM). Three learning models achieved a maximum accuracy of 96.1% for Logistic Regression, 95.3% for KNN, and 95.5% for SVM. For future investigation, similar models will be applied to the broader target genres, such as Pop, Hip-Hop, and Rock.

I. Introduction

With the exponential growth of the music database and large amounts of music in digital formats, Music Information Retrieval (MIR) has become an imperative research area (Silla, 2008). The main task of MIR is the Automatic Music Genre Classification (AMGC), which is to categorize and organize songs based on the genre, according to some characteristics of the music, such as rhythmic structure, harmonic content, and instrumentation (Tzanetakis & Cook, 2002).

The conventional procedure for organizing music content is the manual use of meta information tags, such as ID3 tags, which include song title, album, year, track number, and other specific details of the file content. (Hacker, 2000). The introduction of supervised learning advanced AMGC using the algorithms of K-Nearest Neighbor (KNN), Gaussian Mixture Models (GMM), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), and Artificial Neural Networks (ANN).

The classification of the sub-genres of a specific genre, however, is more challenging to accurately differentiate against each other due to the musical similarity. In this paper, different supervised learning, including state-of-art algorithms, such as KNN and SVM, were compared with the Logistic Regression model to measure the effectiveness of automated sub-genre classifiers.

The focused main genre of this project was RnB, and the subgenre was Neo-Soul. The “I Love My Neo-Soul” playlist by Spotify was the goal playlist (Spotify, n.d.). The RnB genre is a musical genre that was started by the African-American communities in the 1940s, and the music can be described as a “style of music […] that combines elements of pop, gospel, blues and jazz with a strong back beat” (Library Of Congress, n.d.). Some of the popular sub-genres of RnB include Funk, Soul, Neo-Soul, and Alternative RnB, Neo-Soul was chosen as the target because it is an under-valued subgenre that had a massive impact on modern contemporary music.

This project was started as a way to build a songwriting tool for musicians, to provide a way for musicians to check whether their song will fit into a very specific Spotify playlist. In recent years, Spotify official playlists have become a crucial part of gaining new fans and listeners for songwriters and producers. By being placed into a Spotify official playlist, a particular song can potentially have an exponentially higher number of Spotify streams and new listeners. Therefore, being able to use machine learning to predict what the most suitable playlist is for a new song, will be extremely beneficial for songwriters and music producers.

II. Dataset

Since there is no publicly available dataset for specific audio features in the Neo-Soul genre, I collected all the related data by accessing Spotify API. Through this process, I was able to collect audio features such as Acousticness, Danceability, Energy, Instrumentalness, Key, Liveness, Loudness, Mode, Speechiness, Tempo, Time Signature, and Valence for 38 different Official Spotify playlists.

3.1 Data Description

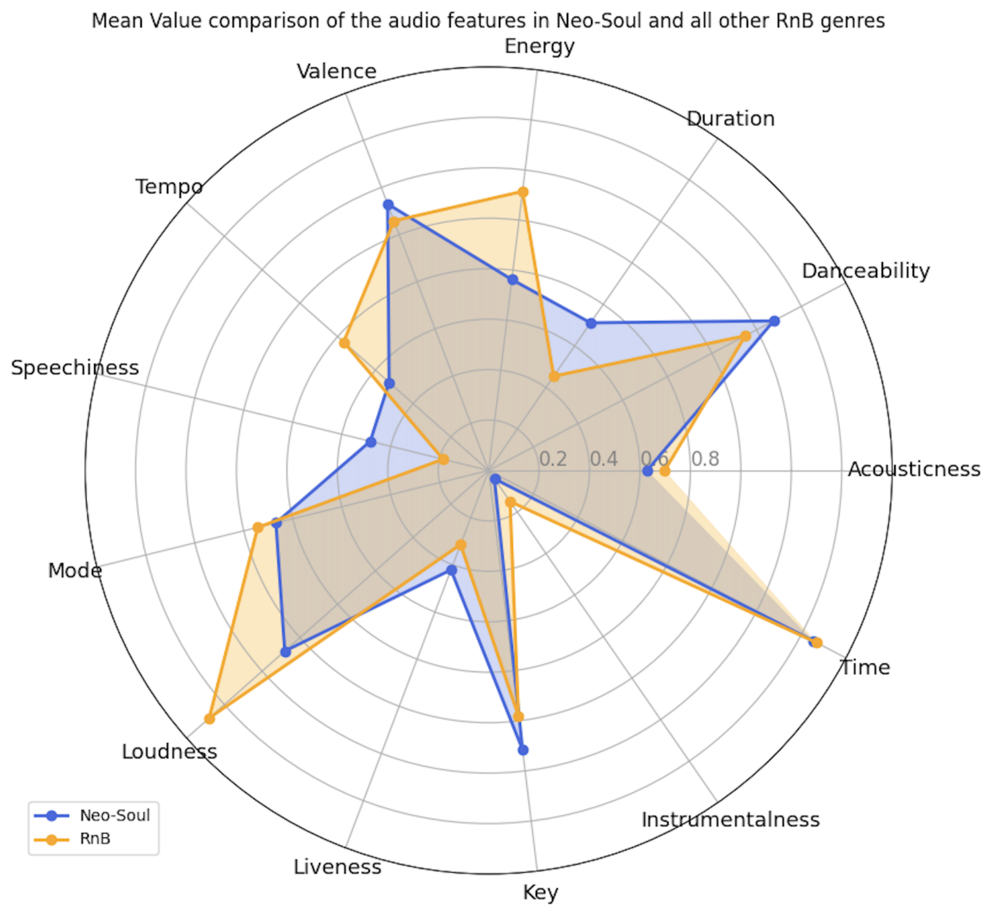

Figure 1. Mean Value comparison of audio features in Neo-Soul and other RnB subgenres.

As seen in Figure 1, the Neo-Soul genre generally is less loud and has a lower energy level than the other RnB genres. On the other hand, the Neo-Soul genre shows higher Danceability and is generally longer in song length than the rest of the genre.

Features like Loudness, Energy, Speechiness, Duration, and Tempo showed a significant difference of over 0.1. Features such as Key, Danceability, Instrumentalness, and Liveness showed a slight difference between 0.05 and 0.07, while Mode, Valence, Acousticness, and Time showed a very little difference under 0.04.

Feature selection was done based on this data, and I removed the Time feature since it showed very little difference between the datasets and used the rest of the features as input for training the model.

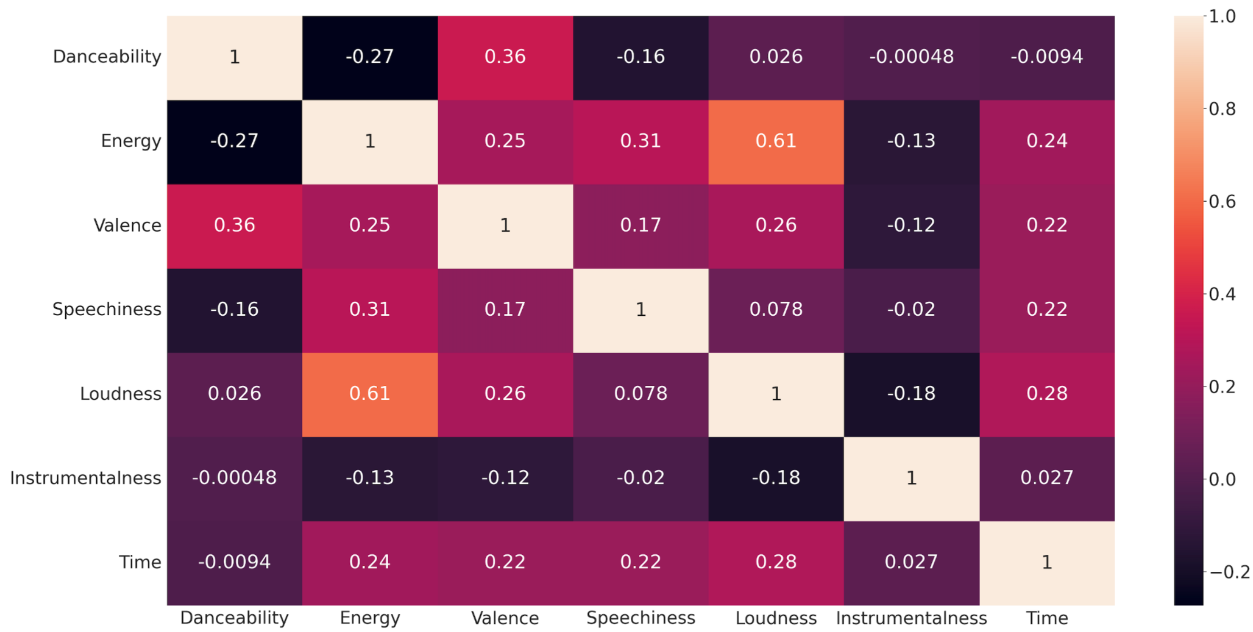

Figure 2. Correlation Matrix of Audio Features in Neo-Soul.

Neo-Soul, as in Figure 2, has distinct characteristics such as high correlation between Energy and Loudness, and Danceability and Valence.

3.2 Data labeling

This project is a two-class problem and the output is either 1 or 0. The target of 1 is assigned for songs that fit into the target playlist: “I Love My Neo-Soul Playlist.” (Spotify, n.d.). The target of 0 is assigned to all the other RnB subgenres, including Funk, New Jack Swing, Soul, Indie RnB, Chill RnB, and Latino RnB.

3.4 Data split

1795 songs were split randomly into 80% training sets (1436 songs) and 20% testing sets (359 songs), using the train test split feature of Scikit-learn. Training sets were also used to tune the model parameters.

III. Experiments

4.1 Performance Metrics

The performance metrics used in this study are:

1. Accuracy: *acc*

These metrics were computed with the Scikit-learn python library.

4.2 Logistic Regression

Logistic regression is a supervised classification system that outputs a discrete outcome given an input variable. The most common logistic regression model outputs a binary outcome and only has two outcomes. There are also multinomial logistic regression models that can have more than two outcomes. Logistic regression is a useful machine learning model for both binomial and multinomial classification problems.

4.3 K-Nearest Neighbor

The K-Nearest Neighbors Classifier, or KNN, classifies a test point by finding the nearest K training points using a given distance formula, and then assigning the majority class of these K training points to the test point. The most significant hyper-parameter of KNN is the positive K, the number of neighbors to be queried for each test point.

4.4 Support Vector Machine

The support vector machine is another machine learning algorithm that classifies a test point by the maximum margin hyperplane separating two classes of data. Non-linear data can be projected to higher dimensional space using the Kernel method (Mandel & Ellis, 2005).

IV. Result and Discussion

The results of the accuracy of each model are summarized in Table 1. Looking at the accuracies, Logistic Regression (LR) had the highest accuracy of 0.961 while both the K-Nearest Neighbor (KNN) and Support Vector Machine (SVM) also produced a very high accuracy score of 0.953 and 0.955. However, the True Positive score came out extremely low compared to the True Negative for all three models, with 340:3 for the LM model. This is probably due to the fact that the sample size of Target 1 (“I Love My Neo Soul Playlist”) is extremely small compared to the sample size of Target 0 (all the other RnB subgenres).

	Accuracy
Logistic Regression	0.961
K-Nearest Neighbor	0.953
Support Vector Machine	0.955

Table 1. The accuracy comparison between three machine learning algorithms: Logistic Regression, K-Nearest Neighbor, and Support Vector Machine. All three displayed high accuracy over 0.9.

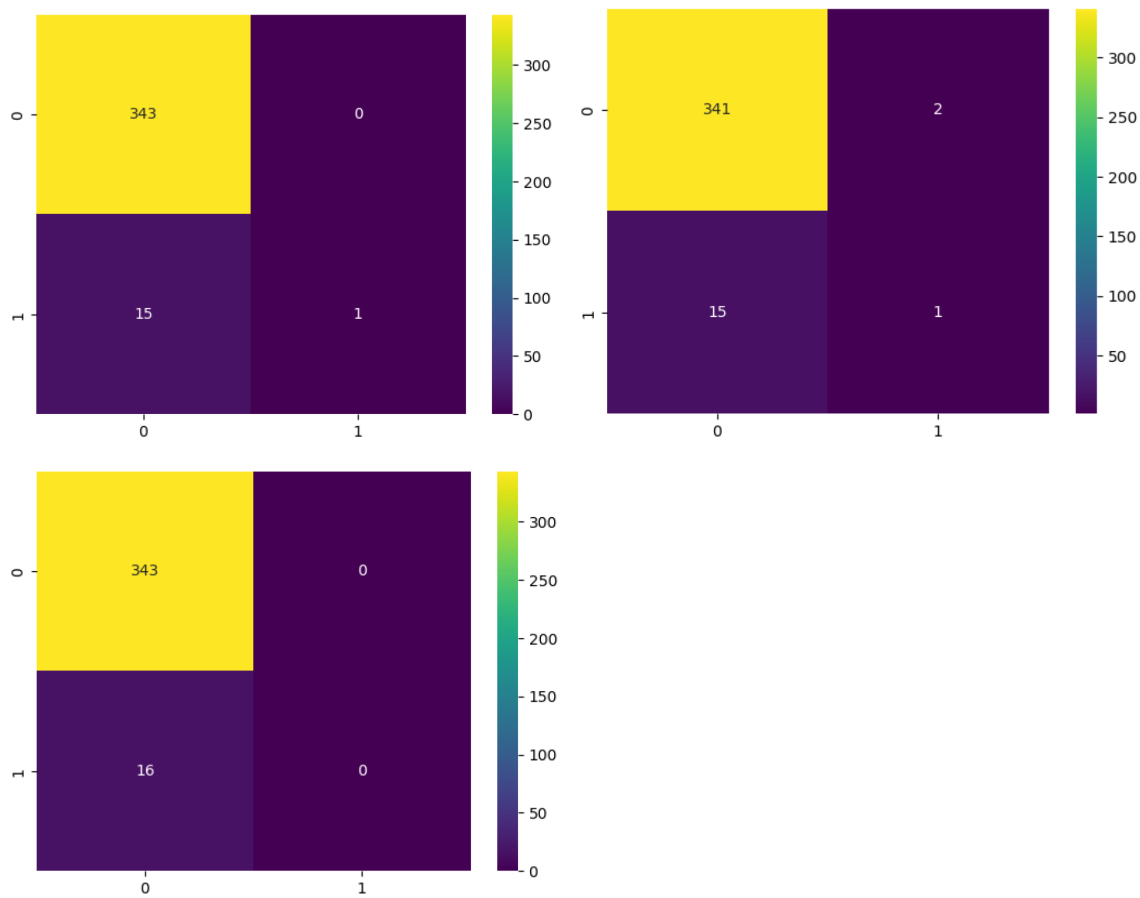

Figure 3. The confusion matrix of Logistic Regression, K-Nearest Neighbor, and Support Vector machine are presented. The yellow represents the true negative. The y-axis represents the true label and the x-axis represents the predicted label. The first column of the first row corresponds to the True negative and the second corresponds to the false positive.

In the testing set of 359 songs, 16 songs were Target 1 and the other 343 songs were Target 0. In Figure 3, all three algorithms displayed high accuracy in predicting the true negative. Out of 343 Target 0 songs, Logistic Regression (LR), K-Nearest Neighbor (KNN), and Support

Vector Machine (SVM) respectively showed the true positives of 343, 341, and 343 songs. LR and SVM showed 0 false positives and KNN showed 2 false positives. Both LR and KNN showed 15 false negatives and 1 true positive, while SVM showed 16 false negatives and 0 true positives. It is important to notice that SVM did not predict any of the songs in the testing set as a Target 1.

V. Conclusion and Future Work

This project was done on a very small subgenre and on a small Spotify Official playlist. In the future, I plan to scale this project into other genres such as Pop, Hip-Hop, and Rock. My goal with this project is to ultimately create a tool where songwriters and producers can check whether their song fits their goal Spotify playlist before releasing their music.

One of the apparent problems was the imbalance of Target 1 songs compared to the Target 0 ones. In the testing set of 359 songs, only 16 songs were considered under the Neo-soul genre (Target 1) and the other 343 songs were under other RnB genres. Therefore, the confusion matrix in Figure 3, shows highly polar results toward the false negative. As the portion of the target is extremely small compared to the entire dataset, it may have given the algorithm an incentive to classify the song into Target 0, which is suggested by 0 false positives and 0 true positives of the confusion matrix of SVM. For future investigation, the polarity can be supplemented with an addition of synthetic data for Target 1, to increase the number of other unpopulated cells of the confusion matrix.

Moreover, the effectiveness of SVM may have to be reconsidered because the linear model was applied without the kernel. Since multiple features of song data were utilized, employing non-linear SVM may yield better accuracy and reduce false negatives.

In terms of the data set, the Spotify API can provide unlimited amounts of data for any of the songs on Spotify. Therefore, using this same methodology, I will be able to apply similar machine-learning techniques to other genres.

Furthermore, other machine learning techniques such as Gaussian Mixture Models (GMM), Linear Discriminant Analysis (LDA), Artificial Neural Networks (ANN), Decision Tree, Random Forest, or AdaBoost Classifier can be used for subgenre classification.

VI. References

Bahuleyan, H. (2018). Music genre classification using machine learning techniques. *arXiv preprint arXiv:1804.01149*.

Basili, R., Serafini, A., & Stellato, A. (2004, October). Classification of musical genre: a machine learning approach. In *ISMIR*.

Clark, S., Park, D., & Guerard, A. (2012). Music genre classification using machine learning techniques.

Ghildiyal, A., Singh, K., & Sharma, S. (2020, November). Music genre classification using machine learning. In *2020 4th international conference on electronics, communication and aerospace technology (ICECA)* (pp. 1368-1372). IEEE.

Hacker, S. (2000). *MP3: The definitive guide*.

Mandel, M. I., & Ellis, D. P. (2005). Song-level features and support vector machines for music classification.

Library Of Congress. *Rhythm and Blues.* (n.d.). Library Of Congress. Retrieved November 22, 2023, from https://www.loc.gov/collections/songs-of-america/articles-and-essays/musical-styles/popular-songs-of-the-day/rhythm-and-blues/#:~:text=The%20term%20%22rhythm%20and%20blues,to%20secular%2C%20not%20religious%20music.

Quinto, R. J. M., Atienza, R. O., & Tiglao, N. M. C. (2017, November). Jazz music sub-genre classification using deep learning. In *TENCON 2017-2017 IEEE Region 10 Conference* (pp. 3111-3116). IEEE.

Silla, C. N., Koerich, A. L., & Kaestner, C. A. (2008). A machine learning approach to automatic music genre classification. *Journal of the Brazilian Computer Society*, *14*, 7-18.

Spotify. *I Love My Neo-Soul*. (n.d.). Spotify. Retrieved November 22, 2023, from https://open.spotify.com/playlist/37i9dQZF1DX8MbMfAHb8U0?si=XdORrUSdSC21B6GzZImerQ&pi=u-PeSfaXrXR9ye&nd=1&dlsi=aaaf4a499b4e454d

Tzanetakis, G., & Cook, P. (2002). Musical genre classification of audio signals. *IEEE Transactions on speech and audio processing*, *10*(5), 293-302.