Music Genre Classification with Machine Learning

Bryan McKenney

May 4, 2022

1 Introduction

Is it possible to determine what genre a song is based on limited data about that song? If a song is uploaded to Spotify without a genre given, in which genre search should it show up? These are the questions that this project hopes to answer. Using Spotify's API, a data set shall be constructed that contains songs from a select number of genres. For each song, the data set will contain the song name, artist name, genre, and features of the audio such as "acousticness," "danceability," "energy," "instrumentalness," "loudness," and "tempo." The goal is to predict the genre of a given song based on its other data, and learning which features matter most for determining genre would be interesting as well, making this an inference problem. Genres are non-numeric labels, so this is a classification problem. It is possible that Spotify's song data will not be enough to accurately predict song genres, but this is a risk that must be taken for curiosity's sake and for the sake of helping people find songs of the genre that they want to listen to (once this completed project is sold to Spotify for millions).

2 Related Work

Kumar et al. (2018) used a data set drawn from the Spotify API as well as lyrics from LyricsFreak to classify songs into one of four genres — Christian, Metal, Country, and Rap. They did this in two different ways: applying multiple machine learning algorithms to preprocessed lyrics, and using a deep convolutional neural network on album covers. Ignatius Moses Setiadi et al. (2020) used a Spotify data set with five genres to determine which audio features matter most to genre classification. They used SVM for classification. Elbir et al. (2018) used the GTZAN music data set, tried multiple machine learning algorithms for genre classification, and had the most success with SVM. Sharma et al. (2018) and Chaudary et al. (2021) both used GTZAN as well and classified music genres with modified versions of SVM (they had ten and five genre classes, respectively).

3 Methodology

The data set will be generated by querying the Spotify API from a Python script and collecting one thousand songs and their audio data from different albums of five distinct genres (Rock, Rap, Pop, Country, and Metal) — the same number of songs for each genre. This data set will then be loaded into an R program for analysis. 20% of the data will be randomly chosen to make up the test set, while the rest will be the training set. Five columns will be added to this training set, one for each genre, with values of 1 and 0 (or -1) so that multi-class classification can be done using the "one vs all" method. The machine learning methods logistic regression, SVM (using the built-in R function which implements the "one vs one" strategy), and LASSO will be applied to these training sets in an attempt to most accurately classify genres in the test set. λ for LASSO will be chosen by cross-validation, and this method will help determine which predictors are most (and least) important. Test set prediction error will be measured by CE (classification error), which should be at most 20% for this project to be successful. The reason for this is that an 80% accuracy rate is sufficient for predicting song genres — it is not critical that the predictions are always correct (if a genre search on Spotify includes a few songs of a different genre, that would not be a big deal).

4 Results

Logistic regression and SVM performed equally well with 59% accuracy rates. LASSO had a 57% accuracy rate. Country, Metal, and Rap proved easy to correctly predict for all three methods, while Pop and Rock were more ambiguous. LASSO's feature selection showed that for predicting Rock, the audio features Instrumentalness, Key, Liveness, Mode, and Tempo were useless. For predicting Rap, the audio features Duration, Key, Tempo, and Time Signature were useless, while the features Explicit, Danceability, Energy, and Speechiness were important. For predicting Pop, the features Liveness, Speechiness, Tempo, and Valence were useless, while the feature Energy was important. For predicting Country, the features Acousticness, Instrumentalness, and Speechiness were important. For predicting Metal, the features Key, Liveness, and Time Signature were useless, while the features Danceability and Energy were important. Overall, Key, Liveness, Duration, Mode, Time Signature, Valence, and Tempo were fairly unhelpful features, while the others were useful for classification. Removing all of the aforementioned unhelpful predictors only reduced the accuracy rate by 0.5% (using SVM).

On similar problems, Kumar et al. (2018) got a best accuracy rate of 84%, Ignatius Moses Setiadi et al. (2020) got a best accuracy rate of 80%, Elbir et al. (2018) got 73%, Sharma et al. (2018) got 87%, and Chaudary et al. (2021) got 94%. Most of these results came from modified versions of SVM. The goal of this project was to reach an 80% accuracy rate, but, given that I was just using basic SVM, a 59% accuracy rate is not terrible — especially because pop songs *are* rap and country songs a lot of the time these days,

which probably accounts for the difficulty in accurately classifying songs labelled as Pop. A music genre search where 41% of resulting songs may sound similar but are not called the same genre is not the end of the world, although it is also not ideal. Using more data, features, and more complex classification methods might improve the results.

References

Chaudary, E., Aziz, S., Khan, M. U., and Gretschmann, P. (2021). Music genre classification using support vector machine and empirical mode decomposition. In *2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC)*, pages 1–5.

Elbir, A., Bilal Çam, H., Emre Iyican, M., Öztürk, B., and Aydin, N. (2018). Music genre classification and recommendation by using machine learning techniques. In *2018 Innovations in Intelligent Systems and Applications Conference (ASYU)*, pages 1–5.

Ignatius Moses Setiadi, D. R., Satriya Rahardwika, D., Rachmawanto, E. H., Atika Sari, C., Susanto, A., Wahyu Mulyono, I. U., Zuni Astuti, E., and Fahmi, A. (2020). Effect of feature selection on the accuracy of music genre classification using svm classifier. In *2020 International Seminar on Application for Technology of Information and Communication (iSemantic)*, pages 7–11.

Kumar, A., Rajpal, A., and Rathore, D. (2018). Genre classification using feature extraction and deep learning techniques. In *2018 10th International Conference on Knowledge and Systems Engineering (KSE)*, pages 175–180.

Sharma, S., Fulzele, P., and Sreedevi, I. (2018). Novel hybrid model for music genre classification based on support vector machine. In *2018 IEEE Symposium on Computer Applications Industrial Electronics (ISCAIE)*, pages 395–400.