Benjamin Lee, Juan Echevarria, Justin Gelabert

Professor Mingon Kang

CS 489

6 May 2021

Music Genre I Write-Up

Our motivation for this project came from our mutual interest in the organizational capabilities of AI. In this case, grouping together different songs based on different genres that Spotify recognizes and how to classify them. Our project has the potential to be used to better organize music into different genres within a listing of songs, similar to the newest update to Spotify which allows the user to select a specific genre from their liked songs. For our project, Juan scraped and created the data set for our project, Justin completed the data preprocessing and preparation to be used as input for the AI models, and Ben implemented the AI models and graphed the results.

We decided to create a multi-class classifier to predict the song genre using the song attributes found in the Audio Features of the Spotify API. These parameters include acousticness, liveness, speechiness, instrumentalness, loudness, danceability, energy, valence, tempo, and key.

We originally were attempting to use the Top Spotify songs from 2010-2019 dataset found on Kaggle[1]. We found that the dataset was too biased towards the "dance pop" genre, and it was far too difficult for the machine learning models to differentiate the subgenres of pop since they were extremely similar to each other. The Kaggle dataset had 50 unique genres in which "dance pop" was the most popular. There was also a bias towards the artists as there were 184

[1] Henrique, Leonardo. "Top Spotify Songs from 2010-2019 - BY YEAR." Kaggle, 26 Dec. 2019, www.kaggle.com/leonardopena/top-spotify-songs-from-20102019-by-year.

unique artists in a dataset that had 603 songs.

To combat this bias, we created a new dataset by scraping songs from the Spotify API using the Spotipy Python library. However, Spotify does not assign genres to individual songs and only gives a list of the genres that the artist is known for. We decided to scrape the top 100 songs from playlists of 10 different genres and assign each song as the genre of its respective playlist. We chose the genre playlists from The Sounds of Spotify[2] and decided to use the following genres: classical, country, jazz, metal, rock, rap, lo-fi, edm, pop, and r&b. Our final dataset had 1000 songs with 988 unique artists that consist of 100 songs of each genre.

When it came to preprocessing our data, we removed the qualitative columns; genre, name, artist, then used min-max scaling from the sklearn library to scale the data to a range from -1 to 1. This was because the values in the dataset had a wide range and the values could be extremely large. When we attempted to fit the data without preprocessing, the program would hang due to the enormous range of numbers. By scaling the data, the models were able to train quickly and efficiently. We used KFold from sklearn to split the data into training and testing sets. We used 10-folds to split the dataset into 10 shuffled parts and used 1 fold for testing and the rest for training.

Since we learned a variety of different classification techniques throughout the course, we decided to implement 6 different machine learning models for multi-class classification. We implemented 3 different support vector machines using the following kernel functions: linear, poly, and rbf. We also implemented K-nearest neighbors using 10 neighbors, logistic regression using the stochastic average gradient descent solver, and a neural network using the multi-layer perceptron classifier. The neural network used the ReLU activation function, adam solver, and was trained with 200 iterations. The learning rate was set at 0.001 and had one hidden layer of

[2] Every Noise at Once, everynoise.com/everynoise1d.cgi?scope=all.

100 hidden units.

We decided to evaluate the overall accuracy of each model as well as the accuracy of classifying each particular genre. This was to see if there was any bias in the data sets or to see if some models were more accurate when classifying particular genres.

Linear SVM had the best overall accuracy with 61.5%. KNN had the worst accuracy with 53.1%. However, most of the accuracies were very close to each other with an only 8.4% difference between Linear SVM and KNN. Classical, lo-fi, and metal were the easiest genres to classify while pop, r&b, and rock were the most difficult to classify. This trend is mostly true amongst all of the models. This made sense since classical, lo-fi, and metal are the most unique in the set of genres we selected. Pop has a lot of overlap with most of the other genres in the set like edm, rap, r&b, rock, etc. However, rock and r&b were unexpectedly difficult to classify, since they seem different enough from the rest of the set. This is likely to be due to the limited number of audio features that spotify assigns to each song.

For reflection, we realized that we could have kept track of the true positive and false positive rates to have a better understanding of how reliable the predictions were rather than only considering accuracy. We could have also explored more with the settings of each model to find optimal hyper parameters. It would also have been helpful to have a larger dataset with more than 100 songs for each genre as the small dataset may have inhibited the training of the neural network. However, while we had room for improvement, the models were able to each successfully classify genres with statistically significant accuracy.