Music Genre Classification

Chi Ung, Omer Diab
Wentworth Institute of Technology

Abstract

Music genre is important for individuals to find more of their favorite music types as well as Music companies to help on make recommendations to their customer. Genre helps on organize a large mass of music so that it's easier to locate and identify a song we don't know. In this project, we developed a Music Genre Classification using Machine Learning techniques. We first created our dataset but extracting music features from 1000 song using LIBROSA library. Then we used some Classification techniques such as K-nearest Neighbor, Support Vector Machine (SVM), and Logistic Regression.

Keywords: Librosa API, SVM, Logistic Regression, MFCC, Chromagram, GTZAN dataset

Introduction

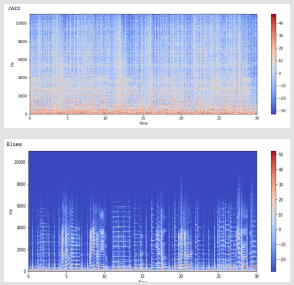
Every musical genre has distinctive features and techniques. Songs that share the same patterns can be grouped in a genre that describe these patterns. These features are a combination of rhythm, speed progression, key, instrumentation, melody, harmony, tempo. When you discover a new song that you like, you often want to hear more songs that sound similar and add to your favorite list of music, but if you don't know the genre, you won't be able to search for that. Also, big companies like YouTube or Spotify need use music genre classification in order to make recommendations to their customer. In this paper we will present our work of applying Machine Learning techniques towards building a classification method in order to identify a genre from music features.

GTZAN Dataset

We used GTZAN genre collection dataset, which provided us 1000 songs with 10 different genres and each genre represented by 100 tracks. Moreover, each song is exactly 30 seconds long and is sampled as 22050HZ Mono 16-bit audio files in .wav format. Even though the dataset has some minor flaws such as collected without the title, copyright and lyrics, which are also very important feature to predict the genres, its dataset itself still made us easily to work on and classifier the training model.

Data processing

We used Librosa Python library to extract features from the .wav files. By using Librosa, it returned the 2D array; for instance, for visualization, LIBROSA provides Spectrogram which is a visual representation of the spectrum of frequencies of sound or other signals as they vary with time. Below are examples of two audios from two different genres:



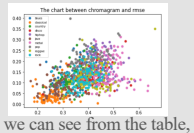
After extracting the audio file to many features, we decided to the features, which were importantly related to the problem we worked on (zero-crossing-rate, spectral centroid, spectral rollof, MFCCs and Chromagram) After that, we scaled all the data by using StandardScaler() and functions it provided.

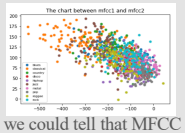
Results

Analysis of MFCCS and Chromagram

Algorithms	Chromagram	MFCC
Logistic Regression	20%	47%
Linear_SVC	19%	41%
RBF_SVC	25%	49%
Poly_SVC	18%	45%
KNN (neighbors = 3)	23%	53%

We used 2 algorithms (Logistic Regression and SVM(and in the case of SVM, we extended to use three different kernels for training the model.





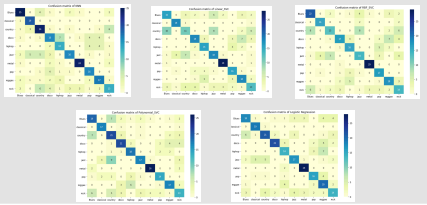
As we can see from the table, we could tell that MFCC has the performance better than Chromagram. The reason could be because of the roles of Chromagram and MFCC. To be specific, according to charts below, because the Chromagram is related to 12 pitch classes, which are very common and widely used in all genres. Due to that reason, it makes the training model and prediction for the genres are very low. On the contrary, since MFCC is very important and it is related to musical instrument, which truly relates to the genres; thus, the predictions and accuracy are higher than Chromagram.

Analysis the dataset

We trained the model with the algorithms, which were very similar when we analyzed the MFCCs and Chromagram (Logistic Regression and SVM).

Models / Algorithms	Training Accuracy	Test Accuracy
Logistic Regression	76%	63%
Linear_SVC	70%	57%
RBF_SVC	79%	59%
Polynomial_SVC	83%	61%
KNN (best k value is 5)	78%	61%

According to the table and confusion matrix graphs we collected, it shows us that the accuracy for all the models are not really bad; however, Linear_SVC has the lowest accuracy and based on the Linear_SVC graph, we realize that for the genre, Blues, it mistakenly predicted with different Country. Because of that, it has an impact to the performance of training model why using Linear_SVC kernel in SVC algorithm. Moreover, not only the genre Blues in most of the algorithms, which they usually mistook with Country genre (8 in RBF_SVC or 6 in Logistic Regression), but also Country genre was mistook with Jazz genre and Polynomial_SVC was affected the most. Overall, the accuracy score for all the algorithms we trained the model is not too bad and we reach up to 83% in Polynomial_SVC algorithm.



Conclusion

To sum up, in order to process with our project, we had to create our own dataset by using GTZAN dataset of 1000 labeled music files, and LIBROSA to extract the music features we need for classification. We used different approaches for the classification of music genres. These approaches are K-Nearest Neighbor, SVM, and Logistic Regression. We found that the best approach was Polynomial_SVC kernel in SVC algorithm with accuracy of 83% in training accuracy and 61% in testing accuracy.

Future work

1.Increase the size of dataset by adding more songs

2.Enhancing the training model to improve accuracy

3.Adding more options such as implementing emotion recognition based on music

4.Develop the mobile application

References

- G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," in IEEE Transactions on Speech and Audio Processing, vol. 10, no. 5, pp. 293-302, July 2002, doi: 10.1109/TSA.2002.800560.
- Sanket Doshi; "Music Feature Extraction in Python". Dec 2018. https://towardsdatascience.com/extract-features-of-music-75a3f9bc265d#:~:text=The%20zero%20crossing%20rate%20is,recognition%20and%20music%20information%20retrieval.
- Deekshith Raya, Vinod Patidar, Yash Khatri, Sai Bhaskar Devatha. "Music Genré Classification". https://www.saibhaskardevatha.co.in/projects/dsp/
- Antonio Jose Homsi Goulart, Rodrigo Capobianco Guido, Carlos Dias Macie. "Exploring different approaches for musicgenre classification". Released: 18 April, 2012. https://www.sciencedirect.com/science/article/pii/S1110866512000151?via%3Dihub