Music Genre Classification using Spectral Analysis and Random Forest Classifier

Hassan Mokari, Fatemeh Akbari, Iman Iraei, Mostafa Abolfazli

Presentation of Machine Learning project

22 December 2022

Introduction

- Music genre classification is the process of using machine learning algorithms to classify music into different genres based on various characteristics. This can be useful for music recommendation systems, music streaming platforms, and other applications where it is important to understand the genre of a particular music. In this research work, the main objective of this project is to make better musical genres prediction as well as comparing predictive performance of different machine learning classifiers.

- GTZAN genre collection dataset with different numbers of audio files and classes. It consists of 1000 audio files each having 30 seconds duration. There are 10 classes (10 music genres) each containing 100 audio tracks.

- Each track is in .wav format. It contains audio files of the following 10 genres such as Blues, Classical, Country, Disco, Hiphop, Jazz, Metal, Pop, Reggae and Rock.

Preprocessing

- **Hyperparameters Optimization:** In our project we use grid search for tunning hyperparameters of each classifier. We have 3 hyperparameters that need to find the optimal values for each of them and they are kernel, regularization parameter and Gamma.

- **Cross Validation:** K-fold Cross-Validation is when the dataset is split into a K number of folds and is used to evaluate the model's ability when given new data. K refers to the number of groups the data sample is split into. Actually, it is used to estimate the performance of machine learning models when making predictions on data not used during training. This procedure can be used both when optimizing the hyperparameters of a model on a dataset, and when comparing and selecting a model for the dataset.

Feature Extraction

The basis of any type of automatic audio analysis system is the extraction of feature vectors. It is noteworthy that the librosa module in Python is used for Feature Vector Extraction which is achieved by applying the following steps.

- Pre-emphasis
- Frame blocking and windowing
- DFT spectrum
- Mel spectrum
- Discrete cosine transform (DCT)
- Dynamic MFCC (Mel Frequency Cepstral Coefficients) features

Machine Learning Models

Once the features are extracted, the following different standard machine learning techniques which are independent of the specific application area will be used as classifiers.

- Decision Tree (DT)
- Random Forest (RF)
- Logistic Regression (LR)
- Support Vector Machines (SVM)
- K-Nearest Neighbors (KNN)
- Gradient Boosting

Evaluation Criteria

Evaluation Criteria: Using evaluation metrics, we can monitor the correctness of the methods and check the model's performance.

- Accuracy
- Precision
- Recall
- F1-Score

Table 1: THE RESULTS OF DIFFERENT CLASSIFIERS ON TESTING DATA

Classifier	Accuracy	F1-Score	Precision	Recall
Random Forest	80.5	80.47	81.08	80.5
Support Vector Machine	77.5	77.75	79.42	77.5
Logestic Regression	76	76.1	76.71	76
K Nearest Neighbors	69	69.16	70.64	69
Gradient Boosting	76	75.62	75.78	76

Confusion Matrix: It is a way to show a summary of the classifier's performance. The confusion matrix shows us the number of correct and incorrect predictions in each class. In this way, we can inform which class is more challenging for our classifier to predict correctly.

As overall results, figure 1 depicts that XGB and RF model had almost identical performance with different iterations and had higher accuracy compared to other models. Furthermore, our results demonstrate that RF and XGB had the highest prediction accuracy (82%), however, LR had the least prediction accuracy (74%) for unseen data.

It is noteworthy that figure 2 depicts the confusion matrix of Random Forest as the best classifier among all denoted classifiers.

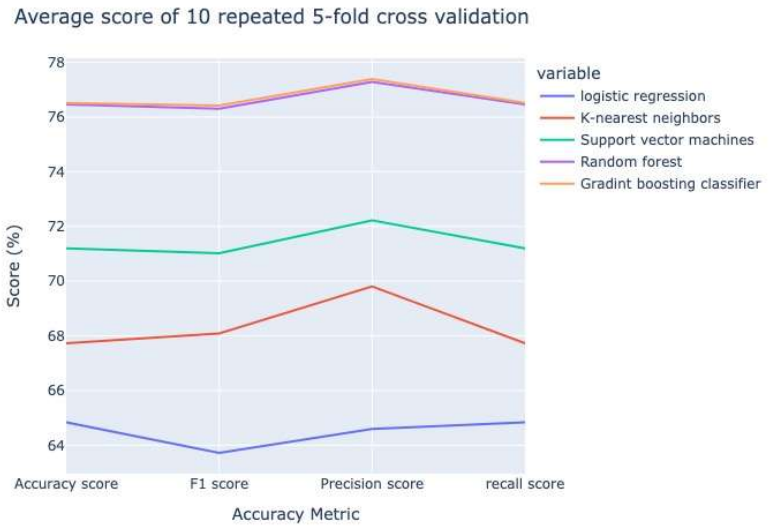

Fig 1: Average scores of five classifiers

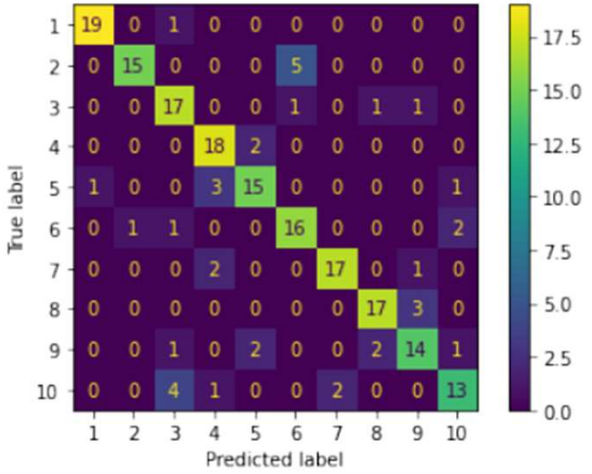

Fig 2: Confusion matrix for Random Forest Classifier

Thanks for your attention.