Predicting Genres for Songs

Ziyad M. Subyani

Abstract

Many of music lovers enjoy listening to it. There are a lot of genres and those genres contains many sub-genres to the point it becomes overwhelming to figure out which genre of music you are listening to, and from here the Idea of this project came. The goal of this project is to use classification models to predict the genres of songs form a wave formatted tracks which is a wave form of an audio. The data was taken from Kaggle which contains a list of tracks translated to numerical data along with its genre. Two models were used in this project Random Forest and XGBoost Classifiers. and I used Seaborn and Matplotlib to Visualize the data and Libraries which is the mother of audio data.

Design

The data is taken form GTZAN dataset which is public dataset for evaluation in machine listening research for music genre recognition. The files were collected in 2000-2001 from a variety of sources including personal CDs, radio, microphone recordings, in order to represent a variety of recording conditions (http://marsyas.info/downloads/datasets.html).

The data consist 9 genres (blues, classical, country, disco, hip-hop, jazz, metal, pop, reggae, rock) for each is assigned to one genre.

Classifying the tracks into one of Theos genres accurately via machine learning models would help listeners to distinguish and understand different genres.

Data

The dataset contains 9990 song each song is 3 seconds long and its in Wave format with 60 features for each, a mean and variance computed over multiple features that can be extracted from the wave format. A few feature highlights include measurements of song tempo(beat per minute) , bandwidth (difference between the upper and lower frequencies), chromogram(intensity of the twelve distinctive pitch classes that are used to study music), and spectral centroid (centre of mass of the spectrum).

Algorithms

1. PCA: Principal component analysis to reduce the features into two components so that it's possible to visualize the data.
2. EDA was used to explore the data and understand it more.
3. Labelencoder was used to encode the labels to be able to see the correlation of data with the label.
4. Splitting the dataset into two datasets for visualization one for the mean and one for the variance.
5. Radar plot was used to compare two songs from different genres.

Models:

The entire training dataset of 9990 was splatted with train-test-split into 80/20 train/test of the data to train the models with. All the scores reported below was calculated after fitting it in Random forest with 100 estimators and max depth of 1000 and 2 random states adding more estimators didn't increase the accuracy and the computing time was longer. The same splits also fitted in the XGBoost with 1000 estimators and 0.05 learning rate.

The matric used to measure the performance of models was the classification rate(accuracy), and confusion matric, and F1 score to see how well the predictions of models on each genre.

Random forest scores:

- Accuracy 0.856
- F1 weighted 0.856
- Precision 0.857
- Recall 0.856

XGBoost scores:

- Accuracy 0.907
- F1 weighted 0.907
- Precision 0.908
- Recall 0.907

Tools

- NumPy and Pandas for data manipulation
- PCA for data compression
- Scikit-learn for modelling
- Matplotlib Seaborn for plotting
- Libraries for visualizing audio files

Communication

In addition to the slides and visuals presented the whole project will be on my Github

Genres