IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Music Genre Classification using Machine Learning

Prof. S. N. UKE[1],Gaurav Singh Negi[2], SatyajitBapusoPatil[3] , SurajYuvraj Lad[4]

1-Assistant Professor, Department of Information Technology, SKN Sinhgad Technical Institute of Technology and Science, Lonavala,Pune,Maharastra,India.

2-4-Student, Department of Information Technology, SKN Sinhgad Technical Institute of Technology and Science, Lonavala,Pune,Maharastra,India.

Abstract:This project was primarily aimed to create an automated system for classification model for music genres. The included steps finding good features that define genre boundaries clearly. There are total of five features, namely MFCC vector, spectral rolloff, spectral centroid,chroma frequencies, zero-crossing rate were used for obtaining features for the classifiers from the GTZAN dataset of genre. Many different classifiers were trained and used to classify, each yielding varying degrees of accuracy in prediction.

***Index Terms* - music, genre, classification, MFCC, chroma, ensemble classifiers, spectral features, GTZAN genre dataset.**

I. INTRODUCTION

"conventional category that identifies pieces of music to a shared tradition or set of conventions called music genre." The "genre" is a subject to interpretation and it is often the case that genres may very blurred in their definition. "Genres" do not have theoreticsoundmusic foundations, e.g. - Indian genres are geographically defined.Despite the lack of a standard criteria for defining genres, the classification of music based on genres is one of the broadest and most widely used. Genre assumes high weight in music caution systems. Genre classification, had been done manually by attaching it to metadata of audio files or including it in album information. This project aims at content-based classification,focusing on information within the audio rather than extraneously attachedinformation.The machine learning approach for classification is used to find suitable features of data, train classifier on feature data, make predictions. The uncommon thing that we have tried is the use of aggregate classifier on fundamentally different classifiers to achieve our end goal.

II. RELATED WORK

The most important work on genre classification using machine learning techniques was developed by Tzanetakis and Cook . The GTZAN dataset was created by them and that is considered as a standard for classificationgenres. Scaringella et al. gives a comprehensive survey of both features and classification techniques used in the genre classification. ChangshengXuetal.have shown how to use support vector machines for this task. Most of the work deals with supervised learning approaches. Riedmiller et al. use unsupervised learning creating a dictionary of features.

III. DATASET

We used the GTZAN dataset from the website of MARYSAS. This is the dataset used. It contains 10 music genres, each has 100 audio clips in .au format. The genres are –Blues, Classical ,Country, Disco, pop, Jazz, Reggae, Rock, Metal. Each audio clips has a length of 30 seconds, are 22050Hz Mono 16-bit files. The dataset incorporates samples from variety of sources like CDs, Radios,Microphone recordings etc. We split the dataset in 0.9 : 0.1 ratio and used 5-fold cross validation for reporting the results.

IV. PRE-PROCESSING

This part contains the conversion of one format of audio files into another, In this preprocessing part involved converting the audio from .au format to .wav format to make it compatible to python's module for reading audio files. The open source Sox utility was used for this conversion.

V. WORKFLOW

To classify our audio clips, we chose 5 features: Mel-Frequency Cepstral Coefficients, Spectral Centroid, Zero Crossing Rate, Chroma Frequencies, Spectral Roll-off. we used different multi class classifiers and ensemble of these to obtain our results.

VI. METHODOLOGY

1. Extraction of Features

5 Features were used to create a single feature vector.

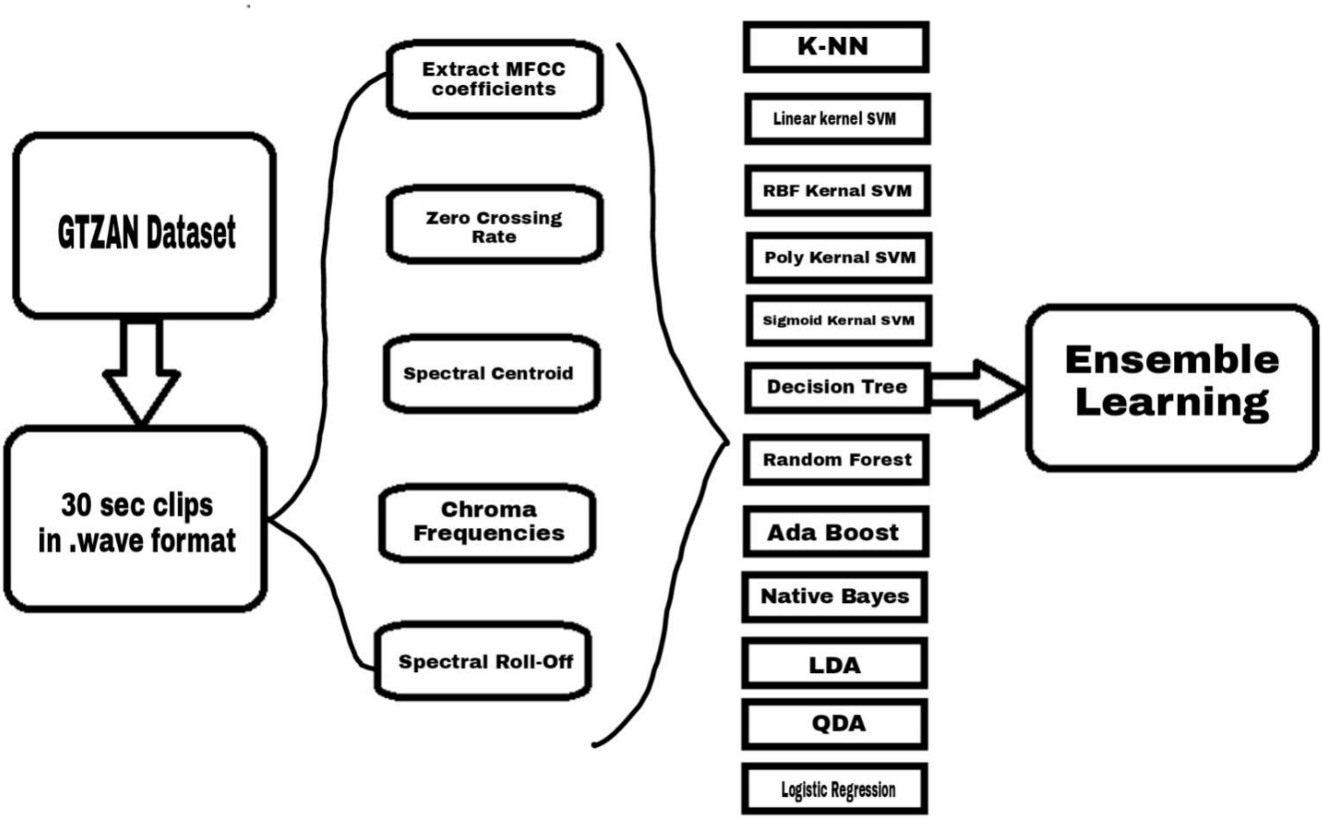

Fig.1. Diagramtic representation of used methods

The **MFCC** feature extraction technique basically includes dormer the signal, applying the DFT, taking the log of the magnitude, and then warping the frequencies on a Mel scale, followed by applying the inverse DCT. Following is the description of various steps involved in MFCC features:

- Dividing the signal into several short frames.
- For each frame, we calculated the periodogram estimate of the power spectrum. This is to know frequencies present in the short frames.
- For each frame, we calculated the periodogram estimate of the power spectrum. This is to know frequencies present in the short frames.

$$M(f) = 1125\ \mathrm{logn}\,(1 + f/700)$$

Fig.2 Formula to work with Mel Scale

- Calculating the logarithm of the filterbank energies It enables humans to have our features closer to what humans can hear.
- We remove the higher coefficients of DCT which can introduce errors by representing changing in the filterbank energies.

The **zero-crossing rate** is the rate of changes of sign along a signal, i.e., the rate at which the signal changes from +ve to 0 to -ve from -ve to 0 to +ve

ZCR=mean(abs(diff(sign(Signal)));

Fig.3 formula to calcculate zero crosing rate.

The **spectral centroid** is a measuredly used in digital signal processing. It describes the "center of mass". actually it is weighted mean of the frequencies present in the sound.
It indicates where the center of mass of the spectrum is located. emotionally, it has a robust connection with the impression of brightness sound.
Consider two songs, one from blues and one from metal. A blues song is generally expected throughout it length while a metal song usually has more frequencies grow towards the end part.

$$\text{Centroid} = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)}$$

Fig.4formula to calculate spectral centroid.

Chroma Frequenciesis a descriptor, which represents the tonal content of a musical audio signal in a liqidized form.. The histogram over the 12-note scale actually is sufficient to describe the chord played in window. It provides a robust way to describe a similarity measure between music pieces.

Spectral rolloff is the frequency below which determine total percentage of spectral energy. Spectral rolloff represents the frequency at which high frequencies drop at 0.

VII. RESULT

Following is the results of current state of music genre classification.

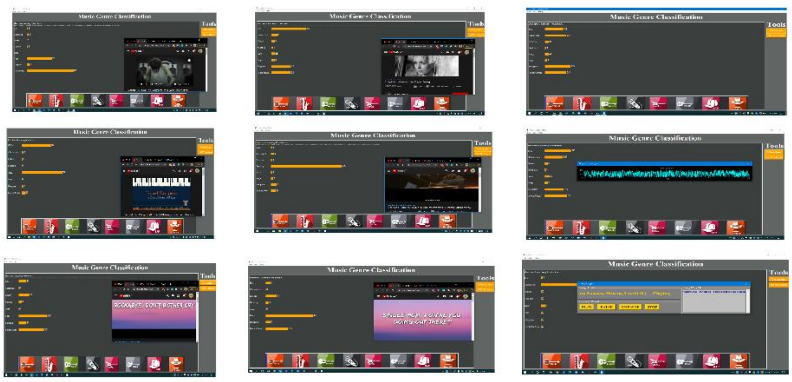

Fig.5 Representss the UI Diagrams.

VIII. CONCLUSION

Themusic genre classification is studied using the Audioset data. The classifier that work best is SVM.We proposed a simple approach to solving the classification problem and we drew comparisons with multiple other complex, robust models. We also compared the models based on the kind of input it was receiving. Some classifier are very efficient for some specific genres like SVM..The highest verified accuracy on the GTZAN dataset is reported at approximately 84%.

IX.REFERENCES

[1] Musical Genre Tag Classification With Curated and Crowdsourced Datasets. Stanford University, Computer Science, 1 edition, 2012.
[2] Bob L. Sturm. Classification accuracy is not enough. Journal of Intelligent Information Systems 41(3):371–406, 2013.
[3] Jan Wu¨lfing and Martin Riedmiller. Unsupervised learning of local features for music classification. In ISMIR, pages 139–144, 2012.
[4]G. Tzanetakis and P. Cook. Musical genre classification of audio signals. IEEE Transactions on Speech and Audio Processing, 10(5):293–302, 2002