International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 04 | April -2020 ISSN: 2582-3930

MACHINE LEARNING AND CHORD BASED GENRE PREDICTION IN MUSIC

I M SHIVAPRASAD
CO-AUTHOR - SUMA S

MASTER OF COMPUTER APPLICATION
DAYANANDA SAGAR COLLEGE OF ENGINEERING
BENGALURU

---*****---

Abstract-In this paper, have advanced a music kind arrangement approach dependent on Mel Frequency Cepstral Coefficients (MFCC). The Mel Frequency Cepstrum (MFC) encodes the force range of a sound. It is determined as the Fourier change of the logarithm of the sign's range. A tweak spectrogram is utilized comparing to the assortment of adjustment spectra of Mel Frequency Cepstral Coefficients (MFCC) will be developed. A spectrogram is a visual depiction of the repeat content in a tune. It shows the power of the frequencies on the y pivot in the predefined time interims on the x hub; that is, the darker the shading, the more grounded the recurrence is in the specific time window of the tune. Every kind existing starting at now is painstakingly structured with a specific number of model vectors which were planned with suitable calculations. A data combination approach containing both element and choice level combination is included for the fitting result. As of late there was an advancement which saw a million-tune dataset being discharged which ready all the tune highlights and its metadata. In any case, we find that our strategy including the notable highlights recorded above improves the precision for recognizing class of a music document.

Index Terms-Modulation Spectrogram, Mel Frequency Cepstrum(MFC),Mel Frequency Cepstral Coefficient(MFCC), Genre Classification.

I. INTRODUCTION

Many companies nowadays use music classification, either to have the option to put proposals to their clients, (for example, Spotify, Soundcloud), or essentially as an item (for instance Shazam). Choosing specific music types is a first step towards this target.

Music sort can be characterized as a classification or rather regular classification that perceives the qualities or attributes of sub-division of the music document having a place with a customary or any traditional set up music structure.
The term Music Genre Classification can be clarified as ordering of music tests. A music kind classifier assumes an indispensable job in declaring melody tests in a primer stage, for example if a new tune has been recorded it will help in sorting the tune into its regular classification. To decide the class of a tune it must be recognized by its interesting sound highlights so its substance can be dissected as for the created wave signals. Another significant perspective is the acknowledgment of the instruments utilized in the melody which is otherwise called the timbral attributes. This assumes an indispensable job in determining the music type dependent on the kind of instrument causes us mean the conventional associate with the music [1].

The Mel Frequency Cepstrum Coefficient (MFCC) for encoding the force range of the sound with the estimation of the Fourier change of the logarithm of the sign's range. Another significant job player here is the spectrogram which encourages us in the visual portrayal recurrence content in a melody. It shows the power of the frequencies on the y pivot contrasted with the time interim on the x hub. This makes the activity simpler in foreseeing the potential kind of the music record consequently satisfying our pursuit.

II. PROPOSED SYSTEM

Music genre classification is not a new problem in machine learning, and many others have attempted to implement algorithms that delve into solving this problem. Using MFCC's has become a popular way to attack this problem and was implemented. Presently, executed the delta and speeding up estimations of the MFCC's also, hence expanding the measure of data being collected from the information.

The Proposed framework utilizes Mel Frequency Cepstral Coefficients (MFCC) and spectrogram, hence to show it further the framework is separated into two stages as follows:

i. Ceps construction phase
ii. Genre classification phase
i. Ceps construction phase- In this phase utilize a python content which causes us to break down and convert each record from the informational index in a portrayal that can be utilized by the classifier and be effectively stored on to the plate. This little

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 04 | April -2020 ISSN: 2582-3930

advance forestalls the classifier to change over the dataset each time the framework is run.

ii. Genre Classification Phase- In this phase a dataset is utilized for taking care of the information in the classifier, which makes a memory model inside itself expressed as relapse model. This procedure is finished by the Logistic Regression module of the scikit-learn library. The python content for this design is expressed. When the model has been made, we can utilize it to anticipate types of other sound documents. For effective further utilization of the produced model, it is for all time serialized to the circle, and is de-serialized when it should be utilized once more. This basic procedure improves execution incredibly. Starting at now, the python content needs to work before any testing with obscure sound record can be performed. When the content is run, it will spare the created model. Once the model has been successfully saved, the classification script need not be run again until some newly labeled training data is available [2].

III. MODEL GENERATION

A dataset is utilized for preparing the classifier, which creates an in-memory relapse model. This procedure is finished by the Logistic Regression module of the scikit-learn library. The python content has been accommodated this reason. When the model has been created, we can utilize it to anticipate sorts of other sound records. For effective further utilization of the produced model, it is for all time serialized to the plate, and is de-serialized when it should be utilized once more. This basic procedure improves execution extraordinarily. Starting at now, the python content must be run before any testing with obscure music should be possible. When the content is run, it will spare the produced model. When the model has been effectively spared, the characterization content need not be run again until some recently marked preparing information is accessible. Not many different strides right now the accompanying:

i. Testing- A python content is utilized for determining the status of new and new sound documents and it helps in de-serializing the recently stored models. In this manner, it marks the new documents.

ii. Output Interpreter- All music documents are arranged and its prepared model is spared to the plate. Additionally, charts are created which are spared in the catalog.

iii. ROC Curves- The Receiver Operating Characteristic Curves are created and spared which means the honesty of the characterized sort after the music record is grouped.

Principal component analysis (PCA) is a measurable system that utilizes a symmetrical change to change over a lot of perceptions of potentially related factors into a lot of estimations of directly uncorrelated factors called head parts (or now and then, head methods of variety). In the wake of doing a PCA on the information we got 90% change and ought to diminish the element measurement.

Dimensionality reduction:Different analysts take measurements, for example, mean change IQR, and so on., to decrease the component measurement. A few analysts model it utilizing multivariate relapse and some fit it to a Gaussian blend model. Here we are taking the mean and upper inclining difference of MFCC coefficients.Support Vector Machine (SVM) is a discriminative classifier formally defined by a separating hyperplane.

IV. METHODOLOGY

Mel-Frequency Cepstral Coefficients MFCC speaks to a lot of transient force range attributes of the sound and have been utilized in the stateof theworkmanship acknowledgment and sound categorization systems. It displays the characters of human voice. This highlight is an enormous piece of the final includes vector. The technique to execute this component is underneath:

– Dividing the sign into a few short edges. The point of this progression is to keep a sound sign consistent.

– For each casing, we determined the periodogram gauge of the force range. This is to realize frequencies present in the short edges.

– Pushing the force spectra into the Mel filterbank and gathering the vitality in each filter to aggregate it. We will at that point know the number of vitalities existing in the different recurrence districts.

– Calculating the logarithm of the filterbank energies in the past It enables individuals to have our features closer to what individuals can hear.

– Calculating the Discrete Cosine Transform (DCT) of the result.

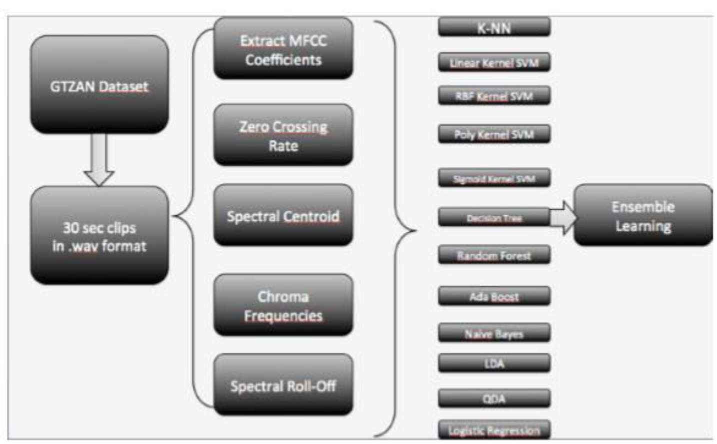

Fig. 1. Diagramtic representation of the method used

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 04 | April -2020 ISSN: 2582-3930

V. CLASSIFICATION

When the element vectors are acquired, we train different classifiers on the preparation set of highlight vectors. Following are the different classifiers that were utilized

– K Nearest Neighbours
– Linear Kernel SVM
– Radial Basis Function (RBF) Kernel SVM
– Polynomial Kernel SVM
– Sigmoid Kernel SVM
– Decision Tree
– Random Forest
– Ada Boost
– Naives Bayes
– Linear Discriminant Analysis (LDA) classifier
– Quadratic Discriminant Analysis (QDA) classifier
– Logic Regression

K-nearest neighbors (KNN): Standard is that the information occurrence of a similar class ought to be nearer in the element space. For a given information point x of obscure class, we can process the separation among x and all the information focuses in the preparation information and appoint the class dictated by k closest purposes of x.

Logistic Regression: Logistic Regression is one of the widely used classification algorithm. This algorithm is used in medical as well as business fields for analytics and classification [3].

Neural Networks: Our second approach was the use of Neural Networks. They are exceptionally inclined to customization, and considering our methodology of a somewhat impossible to miss include vector with a fairly little informational index, this technique appeared to be a decent decision. The thought is to give our system a few sources of info (each information is a solitary element). At that point, we explore through the system by applying loads visit inputs, adding them, and finally utilizing an actuation capacity to acquire the ensuing hubs. This procedure is rehashed on numerous occasions through the shrouded layers, before applying a final initiation capacity to get the yield layer, our prediction.

MFCCs are derived as follows:

1. Take the Fourier transform of (a windowed excerpt of) a signal [4].
2. Map the forces of the range acquired above onto the Mel scale, utilizing triangular covering windows.
3. Take the logs of the forces at each of the Mel frequencies.
4. Take the discrete cosine change of the rundown of Mel log powers, as though it were a sign.
5. The MFCCs are the amplitudes of the subsequent range.

We started by classifying a subset of 6 genres (country, reggae, metal, pop, rock,hip-hop) using an SVM classifier with polynomial kernel. It gave an accuracyof 85%. Anyway, when we attempted stretch out this classifier to 9 classes the precision of classification dropped to 51%. In the following endeavor, we tuned the hyperparameters for the classifiers and made a troupe classifier out of them. Fig shows exactness esteems for different classifiers.

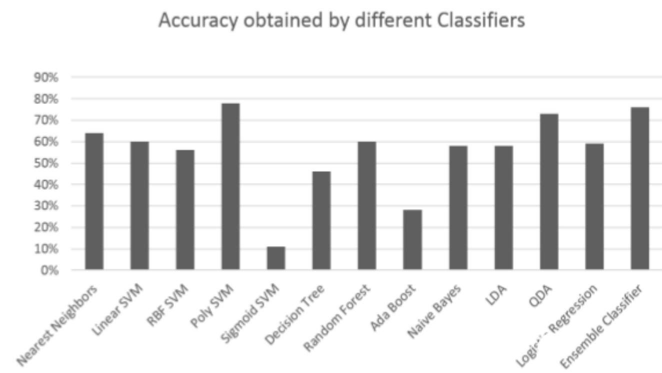

Fig. Accuracy obtained by different classifiers

Classifier	Mean Accuracy	Mean Precision	Mean Recall
K-NN	0.64	0.70	0.64
Linear Kernel SVM	0.60	0.68	0.60
RBF Kernel SVM	0.55	0.78	0.56
Poly Kernel SVM	0.78	0.79	0.78
Decision tree	0.45	0.48	0.46
Random Forest	0.54	0.59	0.54
Adaboost	0.28	0.28	0.28
Naive Bayes	0.57	0.65	0.58
LDA	0.58	0.68	0.58
QDA	0.73	0.77	0.73
Logistic Regression	0.58	0.68	0.59
Ensemble Classifier	0.76	0.81	0.78

Table . Statistics for different classifiers

Classifier	Training Accuracy	Testing Accuracy
K-Nearest Neighbors		53%
Logistic Regression	75.778%	54%
SVM Linear Kernel	99%	52%
SVM RBF Kernel	99%	12%
SVM Poly Kernel	**99%**	**64%**

Table. Accuracy Results

VI. Conclusion

This approach to classify a music genre can be termed as one which needs very few computations and does not even require much of any data. But the result speaks otherwise with respect to the data, it occurs to us that a large quantity of data needs to beprovided and trained from its preliminary stage of the classification process.

We have tried various machine learning algorithms for this project. Our aim is to get maximum accuracy. We have found from our research that we can a get maximum accuracy of 65% by using poly kernel SVM for 10 genre classes. We have also tried to find the best combination of genre classes which will result in maximum accuracy. If we choose 6 genre classes, we were able to get an accuracy of 85%. We chose these labels for the Web Application [classical, hip-hop, jazz, metal, pop and rock] For some songs we can say that it has feature of multiple genres. So, we have also tried to get multiple label outputs based on the probability [5].

VII. REFERENCES

[1]. Music Genre Classification using Machine Learning Techniqueswww.researchgate.net/publication/324218667_Music_Genre_Classification_using_Machine_Learning_Techniques

[2]. Chapter 1.docx | Machine Learning | Statistical Classificationwww.scribd.com/document/416412228/Chapter-1-docx

[3].Mel Frequency Cepstral Coefficients Analysiswww.sites.google.com/site/autosignlan/algorithms-used/mfcc

[5].Music-Genre-Classification/README.md at master http://github.com/DAKSHSEMWAL/Music-Genre-Classification/blob/master/README.md

[6]. Music Genre Classification using Machine Learning Techniques www.groundai.com/project/music-genre-classification-using-machine-learning-techniques/1

[4]. Music Genre Classication http://cs229.stanford.edu/proj2016/report/BurlinCremeLenain-MusicGenreClassification-report.pdf