International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 5 Issue 4, May-June 2021 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Music Genre Classification using Machine Learning

Seethal V[1], Dr. A. Vijayakumar[2]

[1]Master of Computer Application, [2]Professor,
[1,2]Department of Computer Application, Jain Deemed-to-be University, Bangalore, Karnataka, India

ABSTRACT

Music genre classification has been a toughest task in the area of music information retrieval (MIR). Classification of genre can be important to clarify some genuine fascinating issues, such as, making songs references, discovering related songs, finding societies who will like that particular song. The inspiration behind the research is to find the appropriate machine learning algorithm that predict the genres of music utilizing k-nearest neighbor (k-NN) and Support Vector Machine (SVM). GTZAN dataset is the frequently used dataset for the classification music genre. The Mel Frequency cepstral coefficients (MFCC) is utilized to extricate features for the dataset. From results we found that k-NN classifier gave more exact results compared to support vector machine classifier. If the training data is bigger than number of features, k-NN gives better outcomes than SVM. SVM can only identify limited set of patterns. KNN classifier is more powerful for the classification of music genre.

KEYWORDS: *Machine Learning, Mel Frequency Cepstral Coefficients, k-Nearest Neighbors, Support Vector Machine*

How to cite this paper: Seethal V | Dr. A. Vijayakumar "Music Genre Classification using Machine Learning" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-5 | Issue-4, June 2021, pp.829-831, URL: www.ijtsrd.com/papers/ijtsrd41263.pdf

IJTSRD41263

I. INTRODUCTION

A music genre is a conventional category that recognizes a couple of pieces of music as having a spot with a typical practice or set of shows. It is to be recognized from musical design and musical style, yet eventually these terms are a portion of the time used equally. The innovative thought of music infers that these classifications are consistently theoretical and debatable, and a couple of genres may cover.

With the appearance of MP3 and other audio coding schemes, music content investigation has become a developing exploration region. Genres gives an extremely valuable description of a musical piece[12]. The advent of huge music collections has represented the test of how to recover, browse, and suggest their contained items. One approach to facilitate the access of huge music classification is to keep label explanations of all music resources. Labels can be added either manually or automatically. However, because of the high human effort needed for manual labels, the execution of automatic labels is more cost effective. However, an absence of label consistency exists in the music community. This prompts challenges in comparing the performance of genre classification algorithms across databases[3].

In this paper we compared K-NN and SVM classifiers to characterize the following genres: Classical, Jazz, Metal, Pop, Blues, Country, Disco, Hip-jump, Metal, Reggae, Rock. GTZAN dataset is the recommended dataset for the music genre classification. It comprises 1000 sound records each having 30 seconds term. We utilized MFCC to remove data from our information.

RELATED WORKS

Music genre classification has been a hardest task in the field of MIR. Tzanetakis and Cook spearheaded the work on music genre classification utilizing ML techniques. They made the GTZAN dataset as a norm for genre classification.

Nilesh M. Patil.[2] have depicted an automated classification framework model for music genres. It first discovered good features for each music genre. To get feature for the classifiers from the dataset, features like MFCC, chroma frequencies, spectral centroid, zero-crossing rate were utilized. K-NN and SVM classifiers were arranged and used to characterize, each yielding shifting degrees of exactness in prediction.

Sergio Oramas.[3] have shown an way to deal with consider and combine multimodal data representations for genre classification is proposed. Analyses on single and multi-label genre classification are completed, evaluating the impact of the distinctive learned representations and their combinations. Results on the two trials show how the aggregation of learned representations from various modalities improves the exactness of the classification.

Muhammad Asim Ali [4]had is covered ML algorithm that predict the genres of songs utilizing k-NN and SVM. This paper additionally presents similar examination between k-NN and SVM with dimensionality return and afterward without dimensionality decrease by means of principal component analysis (PCA). From results they found that both k-NN and SVM gave more exact outcomes contrast with the outcomes with dimensionality reduction.

II. PROPOSED SYSTEM

The advent of huge music collections has represented the test of how to recover, browse, and suggest their contained items. One approach to facilitate the access of huge music classification is to keep label explanations of all music resources. Labels can be added either manually or automatically. However, because of the high human effort needed for manual labels, the execution of automatic labels is more cost-effective. To solve this issue, we compared K-NN and SVM classifiers to classify the genres by using the GTZAN dataset. From the results, we found that the k-NN gave more accurate outcomes compared to the SVM classifier [3].

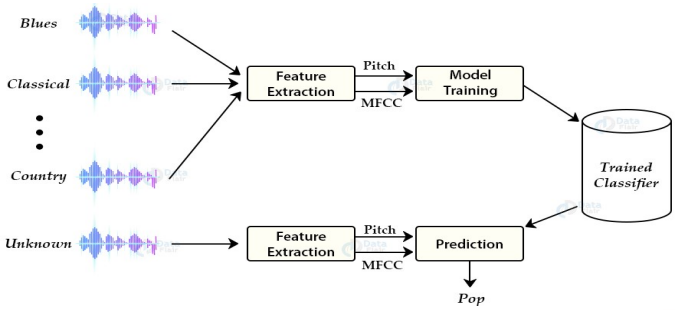

Figure 1 Block Diagram

The first step for genre classification is to separate features and segments from the sound records. It includes recognizing the linguistic contents and disposing of noise. Pitch and MFCC features are removed from sound records. These features are utilized to traina KNN classifier. At that point, new sound records that should be arranged go through a similar feature extraction. The trained KNN classifier predicts which one of the 10 genres is the nearest match [8].

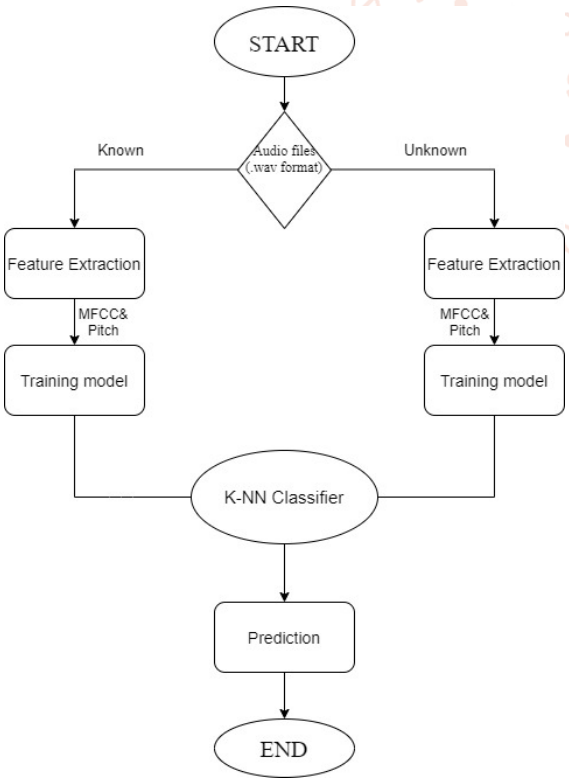

Figure 2 Work Flow

A. Data Gathering

MARSYAS is an open source WWW for sound handling with specific supplement on sound information data users. For music genre classification GTZAN dataset utilized which has a collection of 1000 sound records. Every one of the records is 30 seconds long. Ten classifications are available in this dataset containing 100 tracks each. Each track is in .wav format. It contains sound records of 10 genres [2].

B. Mel Frequency Cepstral Coefficients (MFCC)

The first phase in any music genre classification system is to extract features i.e. recognize the segments of the audio signal that are useful for distinguishing the linguistic content and disposing of the wide range of various stuff which conveys data like background noise, emotion etc. [2].

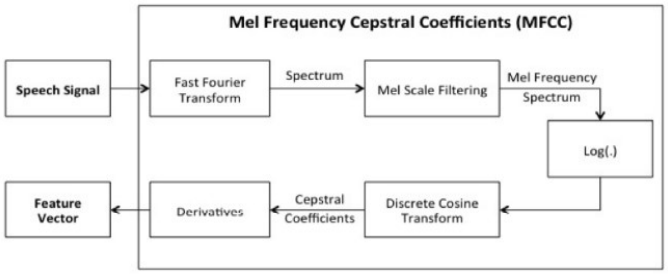

Figure 3 Steps in MFCC

- We partition the signal into a few short frames to keep a n audio signal consistent.
- At that point we attempt to recognize various frequencies present in each frame
- Push the power spectra into the Mel filter bank and assemble the energy in each filter to entire it. With this we get the energy existing in the diverse frequency regions. Equation for Mel scale is:

  $$M(f) = 1125 \ln(1+f/700)$$

- We calculate the logarithm of the filter bank energies. This empowers people to have features closer to what they can hear.
- Calculating the Discrete Cosine Transform (DCT) of the outcome décor relates the filter bank energies with one another.
- We first keep first 13 DCT coefficients disposing of the greater DCT coefficients that can present errors by addressing changes in the filter bank energies[2].

C. k-Nearest Neighbors Algorithm (k-NN)

The principle machine learning technique we utilized was the k-nearest neighbors (k-NN) as it is extremely well-known for its simplicity of execution. K-NN algorithm can be utilized for both classification and regression [9]. These are the two properties that would characterize KNN -

- **Lazy Learning algorithm**- KNN is a Lazy Learning algorithm, it doesn't have a specific training stage and uses all the information for training while classification.
- **Non-parametric learning algorithm**- KNN is a non-parametric learning algorithm, it doesn't expect anything about the secret data[9].

Working of k-NN Algorithm

K-Nearest Neighbors (KNN) Algorithm utilizes feature similarity to predict the values of new data points which further suggests that the new data point will be allocated a value dependent on how eagerly it facilitates with the points in the training set[11].

- For executing any algorithm, we need dataset. So, during the initial step of KNN, we should stack the training similarly as test data.

- After that, we have to take the value of K i.e. the nearest data points. K can be any number.
- Calculate the distance between test data and each row of training data with the assistance of Euclidean distance. To calculate Euclidean distance is.

$$d(\mathbf{p},\mathbf{q}) = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2}$$

- Now, based on the distance value, sort them in ascending order.
- Next, it will pick the top K rows from the arranged array.
- Now, it will allocate a class to the test point dependent on most frequent class of these rows[11].

D. Support vector Machine Algorithm (SVM)

Support Vector Machine is a model for classification and regression. It can deal with linear and non-linear issues and function admirably for some reasonable issues. The algorithm makes a hyperplane which isolates the data into classes[6]. The target of the SVM algorithm is used to settle the decision boundary that can seclude n-dimensional space into classes so it can put the new data point in the right category. This decision boundary is known as a hyperplane. SVM picks the vectors that assistance in making the hyperplane. These extreme cases are called as support vectors, and algorithm is named as Support Vector Machine[6].

III. RESULTS AND DISCUSSIONS

Accuracy of classification by different genres with k-NN and SVM algorithms is varied. The achievement rate of k-NN was 75% yet the blues genre was misconstrued as rock. The k-NN did badly while perceiving blues with a recognizing percentage of 66%. The SVM misconceived classical genre as jazz or hip-hop, yet the rock genre was precisely related to progress rate of 94%. Similarly, the k-NN did well with perceiving entire categories however then again it likewise erroneously distinguished disco with rock and metal with jazz. The k-NN classification had the success rate of country was 73% however with rock genre it was simply 40%. Hip hop genre had the achievement rate of 75%. The classical and raggae had success rate of 69% which is better than SVM however trouble to differentiates with many other genres. By and large we found that k-NN is more successful classifier which gave 70% of accuracy.

Genres	K-NN classifier	SVM classifier
Classical	0.69	0.62
Country	0.73	0.70
Pop	0.70	0.65
Hip-Hop	0.75	0.55
Blues	0.66	0.80
Reggae	0.69	0.14
Accuracy	**75%**	**65%**

IV. CONCLUSION

We have presented a methodology for automatically extracting musical features from audio files and classifying them according to their genre in this paper. We feature extraction and selection, and finally classification process. KNN is a supervised learning classifier that is easy to implement. From results we found that k-Nearest Neighbors classifier gave more accurate outcomes compared to support vector machine classifier. If training data is much larger than number of features, k-NN gives better results than SVM. SVM can only identify limited set of patterns. Overall KNN is more effective classifier which gave 75% accuracy for the classification of music genre.

V. REFERENCES

[1] Wing W. Y. Ng, Weijie Zeng, Ting Wang(2020) "*Multi-Level Local Feature Coding Fusion for Music Genre Recognition*".

[2] Nilesh M. Patil, Dr. Milind U. Nemade(2017) "*Music Genre Classification Using MFCC, K-NN and SVM Classifier*".

[3] Sergio Oramas, Francesco Barbieri, Oriol Nieto, Xavier Serra *(2018) "Multimodal Deep Learning for Music Genre Classification*".

[4] Muhammad AsimAli, Zain Ahmed Siddiqui (2017) "*Automatic Music Genres Classification using Machine Learning*"

[5] Ahmet Elbir; Hilmi Bilal Çam; Mehmet Emre Iyican; BerkayÖztürk; Nizamettin Aydin(2018) "*Music Genre Classification and Recommendation by Using Machine Learning Techniques*".

[6] Anirudh Ghildiyal; Komal Singh; Sachin Sharma (2020) "*Music Genre Classification using Machine Learning*".

[7] KonstantinMarkov; Tomoko Matsui(2014) "*Music Genre and Emotion Recognition Using Gaussian Processes*".

[8] Ali Karatana; Oktay Yildiz(2017) "*Music genre classification with machine learning techniques*".

[9] DivyanshRana(2019) "*Music Recommendation System Using machine learning algorithms*".

[10] J. Umamaheswari; A. Akila(2019) " *An Enhanced Human Speech Emotion Recognition Using Hybrid of PRNN and KNN*".

[11] QhansaDi'Ayu Putri Bayu; Suyanto Suyanto; AndityaArifianto (2019) "*Hierarchical SVM-kNN to Classify Music Emotion*".

[12] JeremyReed, Chin-Hui Lee "A Study on Music Genre Classification Based on Universal Acoustic Models".