Music Genre Classification Using Machine Learning

1M.D.Nevetha, 2A.Nithyasree, 3A.Parveenbanu, 4Mrs.Jetlin CP
1Student, 2Student, 3Student, 4Assistant Professor
Agni College of Technology

***Abstract* - Machine Learning is an application of Artificial Intelligence(AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. In this paper, we've got put forth a expressive style classification approach using Machine Learning technique. Music plays a really important role in people's lives. Music brings like-minded people together and is that the glue that holds communities together. Communities may be recognized by the kind of songs that they compose, or maybe hear. within the area of Music Information Retrieval (MIR), categorizing musical genre could be a challenging task. the aim of our project and research is to seek out a far better machine learning algorithm than the pre-existing models that predicts the genre of songs. Genres may be defined as categorical labels created by humans to spot or characterize the design of music. The concept of automatic expressive style classification has become very talked-about in recent years as a results of the rising of the digital show business. This work presents a comprehensive machine learning approach to the matter of automatic style classification using the audio signal. The system is developed employing a Convolutional Neural Network (CNN) to acknowledge the genres. Here CNN model is trained end to end, to predict the genre label of an audio signal. We are conducting the experiment on the GTZAN data-set, which is widely used public data-set for research in music recognition (MGR).**

***keywords* - Convolutional Neural Network(CNN) , GTZAN data-set , Music Information Retrival (MIR) , Machine Learning(ML).**

I. INTRODUCTION

Music genres are a collection of descriptive keywords that convey high-level information a couple of music clip (jazz, classical, rock...).Music classification is taken into account as a awfully challenging task thanks to selection and extraction of appropriate audio features. Nowadays online music databases growing rabidly and it's very hard for people to accessing those data.one way to arrange and categorize songs is predicated on the genre. Genres are identified by some characteristics of music like rhythmic structure, harmonic content and instrumentation. While unlabelled data is quickly available music tracks with appropriate genre tags is incredibly less Genre classification may be a task that aims to predict genre using the audio signal. having the ability to automatize the task of detecting musical tags allow to form interesting content for the user like music discovery and playlist creations, and for the content provider like music labelling and ordering. Building this technique requires extracting acoustic features that are good estimators of the kind of genres we have an interest, followed by one or multi label classification or in some cases, regression stage. Conventionally, feature extraction relies on a proof processing front-end so as to compute relevant features from time or frequency domain audio representation. The features are then used as input to the machine learning stage. We are visiting make use of GTZAN data-set which is basically famous in Music Information Retrieval (MIR).

II. EXISTING SYSTEM

In existing system, we used k-nearest neighbour (k-NN) to classify the genres. This doesn't give an absolute reasonable correlation between learning strategies for classification of music. It uses filter modelling before Piece wise Gaussian Modelling. However, these improvements don't seem to be statistically significant.This procedure doesn't increases classification accuracy and it doesn't achieve the efficiency prediction. Most of the genre classification studies focuses on finding the simplest set of temporal features, transformations, and filters that best represent the music.The author of the data-set we are using also attempted to search out the set of features that best represent a music.Other studies will try and find combinations of well-known music theories like rhythm analysis to feature new features to the classification problem. We believe that this significantly limit the performance of models because these features are ultimately extracted by humans and that we are going to be missing some important features that would be extracted by a neural network.Other studies have tried to use some AI/Machine learning techniques like Hidden Markov Model to classify music genres, and even SVM. However, they still have limited performance. In recent years, deep learning and neural networks have also been widely applied to classifications problems, including musical style classification. More specifically, using CNN as a music feature extractor was studied by T. LH. Li, A. B. Chan, and A. HW. Chun. They used MFCC audio representation and trained a music pattern extractor to classify style. There are LSTM musical genre classification works being done but mostly focused on lyrics.

III. LITRATURE SURVEY

Music Genre Classification: A N-Gram Based Musicological

Approach Authors: Eve Zheng; Melody Moh; Teng-Sheng Moh, 2017 IEEE 7th International Advance Computing Conference (IACC)

Digitalization of music has grown deep into people's lifestyle. Derived services of digital music, like recommendation systems and similarity test, then become essential for online services and marketing essentials. As a building block of those systems, genre classification is important to support of these services. Previously, researchers mostly focused on low-level features, few of them viewed this problem from a more interpretable way, i.e., a musicological approach. This creates the matter that intermediate stages of the classification process are hardly interpretable, not much of music professionals' domain knowledge was therefore useful within the process. This paper approaches genre classification in a very musicological way. The proposed method takes into consideration the high-level features that have clear musical meanings, so music professionals would find the classification results interpretable. to look at more musicological elements aside from additional statistical information, we use a data-set of only symbolic piano works, including over 200 records of classical, jazz, and ragtime music. Feature extraction and n-gram text classification algorithm are performed. The proposed method proves its concept with experimental results achieving the prediction accuracy averaged above 90%, and with a peak of 98%.

Music Genre Classification and Recommendation by Using Machine Learning TechniquesAuthors: Ahmet Elbir; Hilmi Bilal Çam; Mehmet Emre Iyican; Berkay Öztürk ; Nizamettin Aydin, 2018 Innovations in Intelligent Systems and Applications Conference (ASYU):

Music genre prediction is one amongst the topics that digital music processing is fascinated by. during this study, acoustic features of music are extracted by using digital signal processing techniques and so musical style classification and music recommendations are made by using machine learning methods. additionally, convolutional neural networks, which are deep learning methods, were used for genre classification and music recommendation and performance comparison of the obtained results has been. within the study, GTZAN database has been used and therefore the highest success was obtained with the SVM algorithm.

Genre Recognition Using Residual Neural Networks

Authors; Dipjyoti Bisharad; Rabul Hussain Laskar TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)

Genre is an abstract, yet a characteristic feature of music. Existing works for automatic genre classification compute a group of features from the audio and style a classifier on top of it. Such models, in general, compute these features over a comparatively long duration of the audio. during this paper, a residual neural network based model is proposed for genre classification which is trained on short clips of just 3 seconds duration. Also, traditional genre classification algorithms will assign one genre to an audio clip. However, it's well established that different genres have overlapping characteristics. Considering this ambiguous nature of the genre, the model proposed during this work can assign three genre labels to a music clip, with each genre related to some probability. The proposed model has a blunder rate of 18%, 9%, and 5.5% while predicting into top-1, top-2 and top-3 genres for a music clip respectively. We demonstrate during this work that the predictions made by the classifier align with the broader understood meaning of genre during a realistic setting.

Parallel Recurrent Convolutional Neural Networks Based genre Classification Method for Mobile Devices Authors: Rui Yang ; Lin Feng ; Huibing Wang ; Jianing Yao ; Sen Luo IEEE Access Year: 2020

With the rapid development of the mobile internet of things (IoTs) and mobile sensing devices, an outsized amount of mobile computing- oriented applications have attracted attention both from industry and academia. Deep learning based methods have achieved great success in computing (AI) oriented applications. To advance the event of AI-based Iot systems, effective and efficient algorithms are in urgent need for Iot Edge Computing. Time-series data classification is an on-going problem in applications for mobile devices (e.g. expressive style classification on mobile phones). However, the standard methods require field expertise to extract handcrafted features from the time-series data. Deep learning has been demonstrated to be effective and efficient during this reasonably data. Nevertheless, the prevailing works neglect a number of the sequential relationships found within the time-series data, which are significant for time-series data classification. Considering the aforementioned limitations, we propose a hybrid architecture, named the parallel recurrent convolutional neural network (PRCNN). The PRCNN is an end-to-end training network that mixes feature extraction and time-series data classification in one stage. The parallel CNN and Bi-RNN blocks specialize in extracting the spatial features and temporal frame orders, respectively, and also the outputs of two blocks are fused into one powerful representation of the time-series data.

METHOLOGY

DATA-SET:

The GTZAN genre collection data-set was collected in 2000-2001. 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 subsequent 10 geners:

1.Blues 2.Classical 3.Country 4.Disco 5.Hip-hop 6.Jazz 7.Metal 8.Pop 9.Reggae 10.Rock

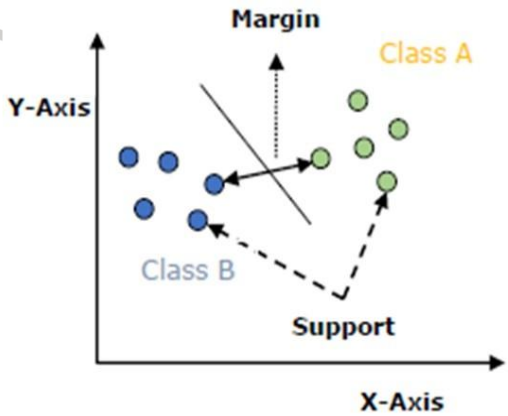
FEATURE EXTRACTION:

Feature extraction may be a process of dimensionality reduction by which an initial set of data is reduced to more manageable groups for processing. A characteristic of those large data sets could be a sizable amount of variables that need lots of computing resources to process. Feature extraction is that the name for methods that select and /or combine variables into features, effectively reducing the quantity of knowledge that has got to be processed, while still accurately and completely describing the first data set. The first step for genre classification project would be to extract features and components from the audio files. It includes identifying the linguistic content and discarding noise.

PREPROCESSED DATA:
We are visiting make use of GTZAN data-set which is de facto famous in Music Information Retrieval (MIR). The data-set comprises 10 genres namely Blues, Classical, Country, Disco, Hip Hop, Jazz, Metal, Pop, Reggae, Rock. Each genre comprises 100 audio files (.wav) of 30 seconds each which means we've got 1000 training examples and if we keep 20% of them for validation then just 800 training examples. we are able to learn the genre of a song or music by paying attention to it for just 4–5 seconds so 30 seconds are little an excessive amount of information for the model to require without delay that's why we decided to separate one audio file into 10 audio files each of three seconds. Now our training examples became tenfold i.e. each genre has 1000 training examples and total training examples are 10,000. So we increased our data-set and this may be helpful for a deep learning model because it always requires more data. it's a process of reworking the raw, complex data into systematic understandable knowledge. It involves the method of sorting out missing and redundant data within the data-set. Thus, this brings uniformity within the data-set. However in our data-set, there was no missing values found meaning that each record was constituted its corresponding feature values.

CONVOLUTIONAL NETWORK MODEL (CNN):
Convolutional Neural Networks (CNNs) are actively used for various music classification tasks like music tag-ging, genre classification and user-item latent feature prediction for recommendation. we apply convolution operation on spectrogram using these filters, then we get four feature maps, as shown in figure. Because different genre will have different of those components, so it's reasonable to use these filters to get high-level feature. CNNs assume features that are in several levels of hierarchy and might be extracted by convolutional kernels. The hierarchical features are learned to attain a given task during supervised training. as an example, learned features from a CNN that's trained for genre classification exhibit low-level features (e.g., onset) to high-level features (e.g., musical instrument patterns).

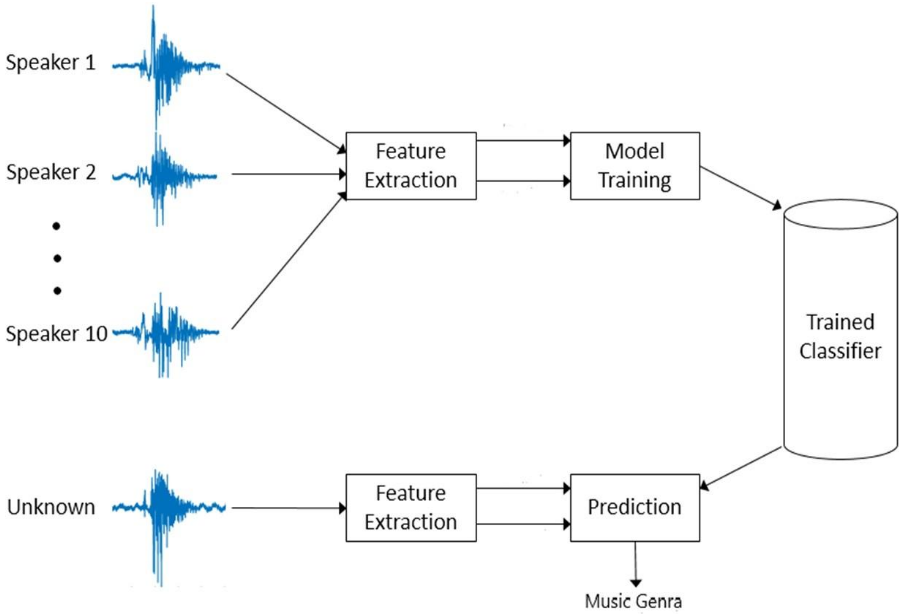
SUPPORT VECTOR MACHINE (SVM):
An SVM model is basically a representation of different classes in a hyperplane in multidimensional space. The hyperplane will be generated in an iterative manner by SVM so that the error can be minimized. The goal of SVM is to divide the data-sets into classes to find a maximum marginal hyperplane (MMH)An SVM model is basically a representation of different classes in a hyperplane in multidimensional space. The hyperplane will be generated in an iterative manner by SVM so that the error can be minimized. The goal of SVM is to divide the data-sets into classes to find a maximum marginal hyperplane (MMH).
The main goal of SVM is to divide the data-sets into classes to find a maximum marginal hyperplane (MMH) and it can be done in the following two steps −
First, SVM will generate hyperplanes iteratively that segregates the classes in best way.
Then, it will choose the hyperplane that separates the classes correctly.



K- NEAREST NEIGHBOUR (KNN):
K-nearest neighbors (KNN) algorithm uses 'feature similarity' to predict the values of new datapoints which further means that the new data point will be assigned a value based on how closely it matches the points in the training set. We can understand its working with the help of following steps −
Step 1 − For implementing any algorithm, we need data-set. So during the first step of KNN, we must load the training as well as test data.
Step 2 − Next, we need to choose the value of K i.e. the nearest data points. K can be any

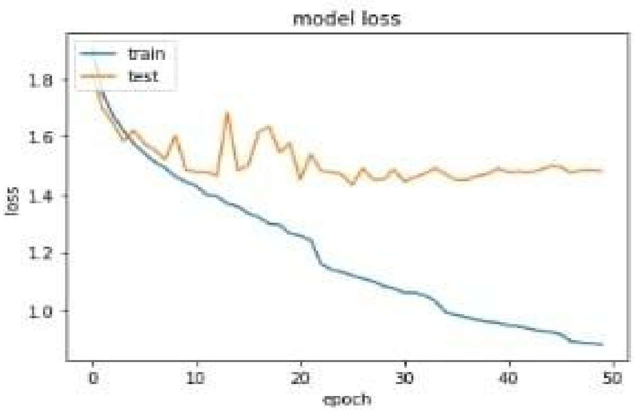
integer. Step 3 − For each point in the test data do the following
3.1 − Calculate the distance between test data and each row of training data with the help of any of the method namely: Euclidean, Manhattan or Hamming distance. The most commonly used method to calculate distance is Euclidean.
3.2 − Now, based on the distance value, sort them in ascending order.
3.3 − Next, it will choose the top K rows from the sorted array.
3.4 − Now, it will assign a class to the test point based on most frequent class of these rows. Step 4 – End

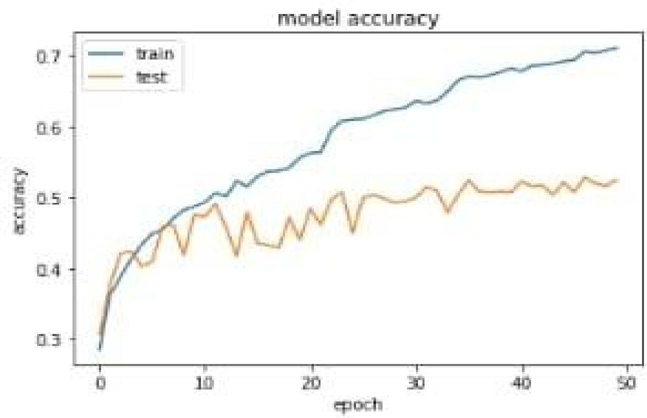
ARCHIECTURE



RESULT AND ANALYSIS

In the previously existing system, the music genre classification system is developed using K-Nearest Neighbour(K-NN) which gives less accuracy . In this classification system, we used K-Nearest Neighbour (K-NN) and Support Vector Machine (SVM) which is developed in Convolutional Kernal with the help of Convolutional Neural Network (CNN) that provide more accuracy compared to the previous system. The testing data-set gives an accuracy of more than 95%.





SUMMARY AND CONCLUSION

We saw how to develop a Convolutional neural network for music genre recognition. In this music genre classification project, we have developed a classifier on audio files to predict its genre. We work through this project on GTZAN music genre classification data-set. It explains how to extract important features from audio files. In this deep learning project we have implemented a K nearest neighbour.

REFERENCE

[1] Music Genre Classification Using Independent Recurrent Neural Network, Wenli Wu ; Fang Han ; Guangxiao Song ; Zhijie Wang, 2018 Chinese Automation Congress (CAC)
[2] Music Genre Classification and Recommendation by Using Machine Learning Techniques, Ahmet Elbir ; Hilmi Bilal Çam ; Mehmet Emre Iyican ; Berkay Öztürk ; Nizamettin Aydin, 2018 Innovations in Intelligent Systems and Applications Conference (ASYU)
[3] Improve Music Genre Classification with Conv Nets, Rafael L. Aguiar ; Yandre M.G. Costa ; Carlos

N. Silla, 2018 International Joint Conference on Neural Networks (IJCNN)
[4] Genre Classification using Word Embeddings and Deep Learning, Akshi Kumar ; Arjun Rajpal ; Dushyant Rathore, 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)
[5] Music Genre Classification: A N-Gram Based Musicological Approach, Eve Zheng ; Melody Moh ; Teng-Sheng Moh, 2017 IEEE 7th International Advance Computing Conference (IACC)
[6] Music Genre Recognition Using Residual Neural Networks, Dipjyoti Bisharad ; Rabul Hussain Laskar TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)
[7] Parallel Recurrent Convolutional Neural Networks Based Music Genre Classification Method for Mobile Devices, Rui Yang ; Lin Feng ; Huibing Wang ; Jianing Yao ; Sen Luo IEEE AccessYear: 2020
[8] On Combining Diverse Models for Lyrics-Based Music Genre Classification ,Caio Luiggy Riyoichi Sawada Ueno ; Diego Furtado Silva, 2019 8th Brazilian Conference on Intelligent Systems (BRACIS)
[9] Automatic Classification of Music Genre Using Masked Conditional Neural Networks, Fady Medhat ; David Chesmore ; John Robinson, 2017 IEEE International Conference on Data Mining (ICDM)