Review Article Impact Factor: 4.226 ISSN: 2319-507X
Manali D. Shirpurkar, IJPRET, 2016; Volume 4 (9): 1244-1250 IJPRET

INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

REVIEW ON MUSIC DATA MINING

MANALI D. SHIRPURKAR[1], PROF. PARNAL P. PAWADE[2]

1. M.E (CSE) 1st Year, P. R. Patil College of Engg, Amravati, Amravati, India.
2. P.R. Patil College of Engg, Amravati, India.

Accepted Date: 15/03/2016; Published Date: 01/05/2016

Abstract*:* The purpose of this research is known about the different features, techniques and classification of Music Data Mining. The features which described in this paper are: Timbre, Intensity, and Rhythm. Also Music Data Mining Techniques are used in this paper such as k-Nearest Neighbors and second one is Cross Validation for mining the music. Music Genre classified into single one which is Machine Learning.

Keywords: Machine Learning, Music Genre classification, k-Nearest Neighbors, Cross Validation.

PAPER-QR CODE

Corresponding Author: MS. MANALI D. SHIRPURKAR

Access Online On:

www.ijpret.com

How to Cite This Article:

Manali D. Shirpurkar, IJPRET, 2016; Volume 4 (9): 1244-1250

INTRODUCTION

Data mining is a process of automatic extraction of novel, useful and understandable patterns from a large collection of data. With the large amount of available data from various sources, music has been a natural application area for data mining techniques. With millions of songs, artists, events touched by potentially billions of listeners online; the resulting online activity opens up vast avenues of research for data science community. [1] While interested in exploring research opportunities involving current Big Data technologies, first attempt to capture current state-of-the-art in music data analysis to clearly identify kinds of datasets, features, analytical techniques that are used by the research community to support various applications and use cases.

Music accounts for a significant chunk of interest among various online activities. As the availability of various online music stores, streaming services, news and podcast services, social networks, and even cloud-based personal music collection. With these developments, moving towards an interesting and contrasting trend. This is reflected by wide array of alternatives offered in music related web/mobile apps, information portals, featuring millions of artists, songs and events attracting user activity at similar scale. Availability of large scale structured and unstructured data has attracted similar level of attention by data science community. Music data mining offers current state-of-the-art in music related analysis. [2] Various approaches involving machine learning, information theory, social network analysis, semantic web and linked open data are represented in the form of taxonomy along with data sources and use cases addressed by the research community.

Music shares a very special relation with human emotions. This often choose to listen to a song or music which best fits our mood at that instant. In spite of this strong correlation, most of the music applications today are devoid of providing the facility of mood-aware playlist generation. Automatic identification of mood in audio songs by utilizing their spectral and temporal audio features. My current work involves analysis of various features in order to learn, train and test the model representing the moods of the audio songs. To produce a system to recognize the mood category of the audio files automatically.

2. Literature Survey and Related Works

Several models have been made to solve this problem like Music Genre Classification with the Million Song Dataset [3], which uses audio features and lyrical features. The Model builds a bag of words for the lyrical features. For the audio features, they used the MFCC (Mel-frequency cepstral cefficients). Their work was unique in how they used lyrical features.

Classification:

Classifying musical genres is a subjective tasks. Robert O. Gjerdingen and David Perrott discovered that participants correctly matched the genre of a song 70 percent of the time after hearing the song for 3 seconds [4]. These results indicate that although there may be a general agreement of what types of genre categories exist, the boundaries separating those genre categories are blurry and unique to each individual. Although humans clearly show some measure of success, machines can assume this role with better accuracy and less effort by creating a feature model of each genre based on certain attributes of each song.

Another paper along the same lines is Automatic Musical Genre Classification Of Audio Signals [5]. A vector of size 9 (mean-Centroid, mean-Roll off, mean-Flux, mean-Zero-Crossings, std-Centroid, std-Rolloff, std-Flux, std-Zero-Crossings, Low-Energy) was used as their Musical-Surface Features vector. Rhythm features were determined and their model was built using both the vectors.

A challenging problem in music data mining is to model expressive dynamics (e.g., variations in tempo and articulation) and discover expression patterns. Recently many data mining techniques, in particular, feature learning techniques have been used to identify useful patterns that are related to the musical expressive performance. In their paper, Grachten and Krebs [6] perform an empirical study to evaluate the utility of several unsupervised feature learning methods in modeling the note intensities of performed music. They use a note centric representation that includes both harmonic and rhythmic characteristics of music contexts and apply different algorithms including non-negative matrix factorization (NMF), principal component analysis (PCA), and Restricted Boltzmann machines (RBM) to learn features. The learned features are evaluated in predicting the note intensities using linear regression models.

The paper by Su *et al.* presents a systematic empirical study that compares BoF variants on three MIR tasks (music genre classification, predominant instrument recognition, and audio tagging) by considering different options on codebook learning and encoding, feature pooling, term weighting, power normalization, and dimension reduction. The study improves the understanding of the BoF model and the analogy between traditional text word sand audio code words, and also provides several interesting insights in applying the model in MIR.

Techniques:

1. k-Nearest Neighbors

The idea behind the k-Nearest Neighbor algorithm is quite straight forward. The algorithm makes use of a similarity function to calculate the similarity between any two sets of data. To classify new song, the system compares the new song with all other songs in the training dataset. It calculates the similarity between the new song and each other song using the provided similarity function.

The system then finds k training songs that have the highest similarity values with the new song. These k songs are referred to as the k nearest neighbors. It then selects the most frequently occurring classes in this set of neighbors and classifies the new song into these classes. The performance of this algorithm greatly depends on two factors, that is, a suitable similarity function and an appropriate value for the parameter k [7].

2. Cross Validation

To minimize the risk of overfitting and to produce more testing data to determine accuracy, the training dataset is partitioned into 4 different subsets to employ a 4-fold cross-validation. The partitioning algorithm randomly picks songs from each category and assigns them to subsets, ensuring to keep a reasonably similar number of songs from each category in each subset. Testing was performed 4 times, each time using one of the subsets as a testing set and the remaining 3 subsets as training sets. Since each song may belong to multiple categories, the number of songs in each subset may differ. However, the representation of each category is approximately equal [7].

Efficient and intelligent music information retrieval (MIR) is a very important topic in music data mining [6]. A key step in MIR is to learn feature representations from available music data. Recently unsupervised feature learning techniques (e.g., sparse coding and deep neural networks) have been used to construct audio code words, leading to the bag-of-frames (BoF) model and a term-document style representation of music. The prime focus is to categorize the audio into different moods. Following are the moods that we have identified currently for our work. Adjacent to the mood category are the adjectives that the mood collectively represents:-

- Happy – (cheerful, funny, romantic, playful)
- Sad – (depressed, frustrated, angry)

- Exciting – (dance, celebration, party)
- Silent – (peaceful, calm)

Songs with similar pattern or their similar audio feature range will be grouped together to yield a particular mood. Hence, a mood based playlist will be provided to the user.

Preprocessing:-

To begin with, to select some songs for each mood based on the survey conducted among different people. To assert the mood of a particular song based on the generalized perspective of human mind.

Feature Extraction:-

JAudio was used for the feature extraction process. Each clip is divided into 0.5 overlapping 32ms-long frames. The extracted features fall into three categories: timbre, intensity, rhythm. These three sets can express mood information to some degree and are very important for mood detection.

1) Timbre features

Happy songs usually sound bright and vibrant, while grief ones sound pensive and gloomy. Timbre features can be used to judge whether the emotion is negative or positive. The timbre features that used are listed as follows: Centroid, Roll off Point, Flux, Zero Crossing, Strongest Frequency via Zero Crossing, Strongest Frequency Via Spectral Centroid, Strongest Frequency Via FFT Maximum, Compactness, MFCC, LPC, Peak based Spectral Smoothness. One has to calculate the mean and standard deviation over all frames.

2) Intensity features

Intensity features can used to judge whether the emotion is very strong or not. For example, if songs express a positive emotion, then using intensity features we can get whether it is enthusiastic or lively. The intensity features are RMS and Fraction of Low Energy Windows[8].

3) Rhythm feature

Through rhythm features, also can get some information about whether the music emotion is positive or negative. Fast songs tend to be happier than slow ones. Extracted rhythm features including Beat Sum, Strongest Beat and Strength of Strongest Beat. Also by calculating the mean and standard, led to 6 rhythm features. The features are extracted and consolidated for each music piece in a standard file format so as to make it easy for mining the relations between these features w.r.t. the corresponding mood of the audio files[8].

Feature Selection:-

There are certain features which give similar values for audio of any mood. Hence such features can hinder the accuracy of the system. After conducting survey and feature extraction process, Information Gain algorithm was used to select the defined features and remove the redundant ones. Information gain helps to determine which attribute in a given set of training feature vectors is most useful for discriminating between the classes to be learned [8]. When a particular classification model has multiple features, there is higher probability that many (if not most) of the features are low information.

REFERENCES

1. Sarla More, and Durgesh Kumar Mishra," Multimedia Data Mining: A Survey", PRATIBHA: INTERNATIONAL JOURNAL OF SCIENCE, SPIRITUALITY, BUSINESS AND TECHNOLOGY (IJSSBT), Vol. 1, No.1, March 2012 ISSN (Print) 2277—7261 March 2012.
2. Shubhanshu Gupta, "Music Data Analysis: A State-of-the-art Survey", Dhirubhai Ambani Institute of Information and Communication Technology, International Journal of Computer Graphics & Animation (IJCGA) Vol.5, No.1, January 2015 18 Nov 2014.
3. Ajay Prasadh Viswanathan, Sriram Sundaraj, "Music Genre Classification", Published: International Journal Of Engineering And Computer Science ISSN: 2319-7242 Volume 4 Issue 10 Oct 2015.
4. Sam Clark, Danny Park, Adrien Guerard, "Music Genre Classification Using Machine Learning Techniques", International MultiConference of Engineers and Computer Scientists 2015 Vol I, IMECS 2015, March 18 - 20, 2015, Hong Kong
5. Aziz Nasridinov1 and Young-Ho Park, " A Study on Music Genre Recognition and Classification Techniques", International Journal of Multimedia and Ubiquitous EngineeringVol.9,No.4a(2014),pp.31aahttp://dx.doi.org/10.14257/ijmue, 2014.
6. Tao Li, Mitsunori Ogihara, George Tzanetakis, "Guest Editorial: Special Section on Music Data Mining", IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 16, NO. 5, AUGUST 2014.
7. Adit Jamdar, Jessica Abraham, Karishma Khanna and Rahul Dubey, "EMOTION ANALYSIS OF SONGS BASED ON LYRICAL AND AUDIO FEATURES", International Journal of Artificial Intelligence & Applications (IJAIA) Vol. 6, No. 3, May 2015

8. Amey Ujlambkar1, Omkar Upadhye2, Akshay Deshpande3, Gunwant Suryawanshi4*,* " Mood Based Music Categorization System for Bollywood Music", International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970) Volume-4 Number-1 Issue-14 March-2014.