Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 98 (2016) 125 – 131

Procedia Computer Science

The 7th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2016)

A Study On Feature Selection And Classification Techniques Of Indian Music

Prafulla Kalapatapu[1,*], Srihita Goli[1], Prasanna Arthum[1], Aruna Malapati[1]

BITS,Pilani University, Hyderabad Campus, Hyderabad-500078,India

Abstract

In this paper we present the effect of four feature selection algorithms namely genetic algorithm, Forward feature selection, information gain and correlation based on four different classifiers (Decision tree C4.5, K-Nearest neighbors, neural network and support vector machine). The feature sets used in this paper are extracted features from the preprocessed songs using MIR Toolbox in MATLAB, which encompass rhythm based, timbre based, pitch based, tonality based and dynamic features. Feature vectors are extracted from music segments from first 30 seconds and last thirty seconds of the music signal (time-decomposition).Experiments were carried out on the three dominant genres of Indian music : Carnatic, Hindustani and Bollywood. Our dataset is small with 290 songs, trimmed to extract the first and the last 30 second percepts. As pure Carnatic and Hindustani music being more prevalent in traditional settings, have limited work done to make their digital copies available but the collection of music we have used consists of songs of some of the most profound singers contributing to each of these genres.For high-dimensional feature sets, the feature selection provides a compact but discriminative feature subset which has an interesting trade-off between classification accuracy and computational effort. The experimental results have shown that the common features selected by each of the feature selection algorithms with respect to classifiers and percentage of classification accuracies for all the classification algorithms. Furthermore, it can be observed from our experiment that information gain based feature selection gives better and consistent accuracies than other feature selection algorithms and Neural network and SVM classifiers are the best suited classifiers for Indian Song dataset.

Peer-review under responsibility of the Program Chairs

Keywords: Music classification; Feature Extraction; Feature Selection; Music information retrieval.

1. Introduction

Rapid growth of digital technologies, the internet, and the multimedia industry has provoked a huge information overload and a necessity of effective information filtering systems and in particular recommendation systems. In the case of digital music industry, current major internet stores contain millions of tracks, which complicate search, retrieval and discovery of music relevant for a user[1]. The features of music can be divided into three categories namely, low level, middle level and high level features. These features are used for the genre classification of songs[2],

* Prafulla Kalapatapu. Tel.: +91-9492922393.
E-mail address: prafulla@hyderabad.bits-pilani.ac.in

Peer-review under responsibility of the Program Chairs
doi:10.1016/j.procs.2016.09.020

music recommendations, classification based on textual features etc., usage of any one of these categories might not give best results[3]. Hence optimal set includes features from all the three categories based on the features variance with time, cost and waveform. The limited scope of the range of features selected through domain knowledge can be widened using the feature selection algorithms in machine learning. Here we propose a comparative study to select this optimal set of features for Indian music.

This paper presents the effect of four feature selection techniques on the classification accuracy of four different classifiers. After features are extracted from the preprocessed songs using MIR toolbox, their effectiveness are measured by comparing accuracies of four traditional classification algorithms applied to only the commonly selected features[4]. The four classification algorithms used in this study are Decision tree C4.5, k-nearest neighbor (kNN), neural network, and Support Vector Machines (SVM) and the four feature selection algorithms used for experiment are Genetic algorithm, Forward feature selection, information gain, correlation based.

The organization of this paper is as follows. The background of the work such as feature extraction, feature selection methods are outlined in Section 2. The four classification algorithms are briefly described in Section 3. Section 4 contains details of dataset we used. Our analysis on the results, including the significance of the feature selection methods, are presented in Section 5. Then we provide some conclusions and future work in Section 6.

2. Background

2.1. Feature Extraction

Feature extraction is a process where a segment of an audio is characterized into a compact numerical representation. In our work features are extracted from the preprocessed songs using MIR Toolbox in Matlab. The feature values are represented in the form of matrices and cells in Matlab. The MIRtoolbox is a collection of Matlab functions for extracting audio features such as tonality, rhythm, and pitch from audio files. The toolbox employs a modular framework which includes preprocessing, classification and clustering functionality along with audio similarity and distance metrics as part of the toolbox functionality. Algorithms are fragmented allowing detailed control with simple syntax, but often suffers from standard Matlab memory management limitations[5]. Because many feature extraction processes share the same initial computations, a range of building block functions are included to avoid running the same calculations multiple times. In this paper rhythm based, timbre based, pitch based, tonality based and dynamic features are extracted.

Rhythm based features include event density, peaks and pulse clarity which capture the rhythmic fluctuations along the audio signal. Timbre based features include segment-wise minimum and maximum of attack time and attack slope, number of zero crossings, rolloff, and brightness, centroid, spread, skewness, kurtosis, flatness, entropy and Mel Frequency Cepstral Coefficients(MFCC)[6]. Pitch based features include pitch and inharmonicity. Tonality based features include chromagram, key, mode, key strength and tonal centroid. Dynamic features extracted are RMS energy and low energy. While some of the features like RMS energy, centroid, zero crossings are uni-dimensional, some features like MFCCs, chromagram, tonal centroid are multi-dimensional. All these 26 features are listed in Table 1 extracted over two segments of each of the songs sum up to a total of 120 dimensions.

2.2. Feature Selection

It is the process of selecting the predominant features from the data set and remove the features that are irrelevant with respect to the task that is to be performed. Feature selection can be extremely useful in reducing the dimensionality of the data to be processed by the classifier, reducing execution time and improving predictive accuracy. Feature selection is preferable to feature transformation when the original units and meaning of features are important and the modeling goal is to identify an influential subset[7]. When categorical features are present, and numerical transformations are inappropriate, feature selection becomes the primary means of dimension reduction. Reducing the dimensionality of the data reduces the computational complexity for bigger datasets such as music data and thus results in faster execution time.

In general, feature selection algorithms can be broadly classified into filter based and wrapper based algorithms. Our proposed work uses two wrapper based approaches: forward feature selection and genetic algorithm and two

Table 1. Features extracted using MIR toolbox

S.No	Features
1	zero crossings
2	tonal centroid
3	key strength
4	Spread
5	Skewness
6	Rolloff
7	RMS Energy
8	Pitch
9	Peaks
10	Mode
11	minimum attack time
12	minimum attack slope
13	MFCCs
14	maximum attack time
15	maximum attack slope
16	low energy
17	Kurtosis
18	Key
19	Inharmonicity
20	Flatness
21	event density
22	Entropy
23	pulse clarity
24	chromagram
25	centroid
26	brightness

filter based approaches with best first strategy: information gain based feature selection and correlation based feature selection were implemented.

2.2.1. Wrapper based algorithms

Wrapper methods are so called because they wrap a classifier up in a feature selection algorithm[8]. Typically: a set of features is chosen; the efficacy of this set is determined; some perturbation is made to change the original set and the efficacy of the new set is evaluated. The problem with this approach is that feature space is vast and looking at every possible combination would take a large amount of time and computation. This means that some heuristic search methods must be developed to find optimum sets of features.

2.2.2. Filter based algorithms

Filter methods apply some ranking over features. The ranking denotes how 'useful' each feature is likely to be for classification. The objective function evaluates feature subsets by their information content, typically interclass distance, statistical dependence or information-theoretic measures. In this study the ranking is computed using information gain and correlation. Once this ranking has been computed, a feature set composing of the best N features is created.

The features selected through each of these approaches are used to classify the data set into the three genres and the resulting accuracies are compared with each other and also with those observed when no feature selection is done.

3. Classification

After selecting the most discriminatory features, we apply k-NN, C4.5, NB, SVM, and PCL to obtain error rates on our testing samples. The classification results of these algorithms are then used to compare the effectiveness of various feature selection methods.

k-NN is a typical instance-based prediction model. By k-NN, the class label of a new testing sample is decided by the majority class of its k closest neighbors based on their Euclidean distance. This is based on learning by analogy, that is by comparing a given test point with training points which are similar to it. When given an unknown point, a k-nearest neighbor (k-NN) classifier searches the pattern space for the k training points which are closest to the unknown point. These k training points are the k-nearest neighbors of the unknown point. The Euclidean distance between two points X_1 and X_2 is obtained by $dist(X_1, X_2) = \sqrt{\sum_{i=1}^{n}(X_{1i} - X_{2i})^2}$, where i is index from 1 to n. In our experiments, k is set as 3[9].

C4.5 is a widely used decision tree based classifier. The implementation of C4.5 in this paper is based on its Revision 8, which was the last public version before it was commercialized. In our experiments, pruned trees and subtree raising techniques are used. In brief algorithm can be summarized in three parts: Selection used to partition training data, Termination condition determines when to stop partitioning and Pruning algorithm attempts to prevent overfitting.[10].

Multilayer perceptron neural network classifier (MLP) model that maps sets of input data onto a set of appropriate outputs. An MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. This supervised learning uses backpropagation momentum algorithm. Learning occurs in the perceptron by changing connection weights after each piece of data is processed, based on the amount of error in the output compared to the expected result. We represent the error in output node p in the nth data point (training example) by $e_p(n) = t_p(n) - v_p(n)$, where t is the target value and v is the value produced by the perceptron. We then made adjustments in the weights of nodes which minimize the error in the entire output. It is among the most practical approaches to certain types of learning problems[11].

SVMs are a kind of blend of linear modeling and instance-based learning. A SVM selects a small number of critical boundary samples from each class and builds a linear discriminant function that separates them as widely as possible. In the case that no linear separation is possible, the technique of kernel will be used to automatically inject the training samples into a higher-dimensional space, and to learn a separator in that space. The SVM used in this paper is a version that implements a sequential minimal optimization algorithm using polynomial kernels. Transforming the output of SVM into probabilities is conducted by a standard sigmoid function[12].

4. Datasets

The dataset chosen for this study were three dominant genres of Indian music, Carnatic, Hindustani and Bollywood. Relatively small dataset is used as pure Carnatic and Hindustani music being more prevalent in traditional settings, have limited work done to make their digital copies available. Due to the lack of openly available datasets, the collection of music we have used consists of songs of some of the most profound singers contributing to each of these genres. The Carnatic dataset consists of a total of 70 songs composed by Anayampatti S. Dhandapani, Dr.N.Ramani, KunnakudiVaidyanathan, LalgudiJayaraman, M.S. Gopalakrishnan, Dr. M. Lalitha and N. Nandini. The Hindustani dataset consists of a total of 120 songs by Ravi Shankar, Anoushka Shankar, Amjad Ali Khan, UstadBismillah Khan, Zakhir Hussain and Kishore Kumar. Unlike classical music, Bollywood music is very popular. Hundreds of Bollywood songs come out every year. There has been a great evolution in Bollywood music over decades. The dataset of the other two genres being smaller, only 100 Bollywood songs are picked uniformly from different time eras. All these songs are collected in new mp3 format. This data was kept constant to facilitate comparison of results.

Musical files for this experiment were obtained from the personal collections of audio CDs from many individuals of the University of BITS, Pilani Hyderabad Campus. The dataset became available in both digital and analog format. Quite a number of musical data for these genres were in analog format and were digitized manually. All of the digital music files (.mp3) were then converted into wav files; the only audio format supported by the existing feature

Table 2. Features selected by each of feature selection algorithm; common features.

Features	Feature Dimension	**Forward Feature Selection**	**Genetic Algorithm**	**Information Gain based feature selection**	**Correlation based feature selection**	**Commonly Selected Features**
zero crossings		Y	Y			
	1	Y	Y		Y	
	2	Y			Y	
tonal centroid	3	Y	Y		Y	
	4	Y	Y			
	5		Y		Y	
	6	Y	Y		Y	
key strength		Y	Y			
spread		Y	Y	Y	Y	Y
skewness		Y	Y	Y	Y	Y
rolloff		Y	Y	Y	Y	Y
RMS Energy		Y	Y	Y	Y	Y
pitch			Y	Y		
	1	Y		Y	Y	
	2			Y	Y	
peaks	3	Y	Y	Y	Y	Y
	4		Y	Y	Y	
	5	Y	Y	Y	Y	Y
	6	Y	Y	Y	Y	Y
mode					Y	
minimum attack time			Y	Y	Y	
minimum attack slope		Y				
	1	Y	Y		Y	
	2	Y	Y	Y		
	3	Y	Y	Y	Y	Y
	4	Y	Y			
	5	Y	Y		Y	
	6	Y				
MFCCs	7	Y			Y	
	8	Y	Y	Y		
	9	Y				
	10		Y	Y		
	11	Y	Y			
	12	Y	Y	Y		
	13	Y	Y			
maximum attack time		Y	Y	Y	Y	Y
maximum attack slope		Y	Y			
low energy			Y		Y	
kurtosis		Y	Y	Y	Y	Y
key			Y			
inharmonicity		Y	Y	Y		
flatness		Y	Y	Y		
event density		Y	Y			
entropy		Y	Y	Y		
pulse clarity			Y		Y	
	1		Y	Y		
	2	Y			Y	
	3	Y			Y	
	4			Y	Y	
	5	Y	Y	Y		
	6		Y	Y		
chromagram	7	Y	Y		Y	
	8		Y			
	9	Y		Y		
	10	Y				
	11		Y	Y	Y	
	12		Y			
centroid		Y	Y	Y	Y	Y
brightness		Y	Y			

Table 3. Accuracies achieved by feature selection algorithm vs classifiers.

	Decision Tree Learning	K-Nearest Neighbors	Neural Network	Support Vector Machines
No Feature Selection	81.7241	86.2069	97.7273	88.6
Genetic Algorithm	82.4138	85.8621	95.4545	88.6207
Forward Feature Selection	87.2414	84.8276	100	89.3103
Information Gain based Feature Selection	86.2069	91.0345	97.7273	90.8
Correlation based Feature Selection	84.4828	86.2068	97.7273	91.0345

extraction tool used at the time of study. The whole dataset was later trimmed to extract the first and the last 30 second percepts by executing certain audio commands through batch processing before extraction began.

5. Results

Table 2 summarizes the common features selected by each of the feature selection algorithms for all the classification algorithms. However there are only 11 features that are commonly selected by all the feature selection algorithms. Table 3 shows the percentage of classification accuracies for all the classification algorithms for each of the feature selection algorithms and as well it can be observed that highest accuracies are obtained with both k-NN and neural network learning using information gain based feature selection. Neural networks performed the best for all the feature selection algorithms, the next best being Support Vector Machines which performed best for correlation based and information gain based feature selection. Neural networks with forward feature selection algorithm gives the maximum accuracy of 100%. It can also be observed that feature selection does not account to much increase in accuracy in case of neural networks but affects decision tree learning the most. Significant observations can be neural network and SVM classifiers performs better and Information gain based feature selection algorithm in combination with all classifiers taken , performed consistently good with minimum 86.2% accuracy.

6. Conclusion and Future Work

In this paper we have reported the effect of feature selection on the accuracy of genre classification on Indian music for three genres under study. Our experimental results prove that feature selection does not always improve the classification accuracy but can still improve the classification accuracy under few circumstances. Hence one must employ proper evaluation methods to understand the effects of feature selection and also the selection of the right classifier. Two most significant observations of this study are that information gain based feature selection gives better and consistent accuracies than other feature selection algorithms and neural network and SVM classifiers are the best suited classifiers for Indian Song dataset.Future work will include further experiments to investigate these findings on improved Indian musical genre classification with bigger dataset.

References

1. Dannenberg, R., Foote, J., Tzanetakis, G., Weare, C.. Panel: new directions in music information retrieval. In: *Proc. Int. Computer Music Conf.(ICMC), Habana, Cuba*. 2001, .

2. Tzanetakis, G., Cook, P.. Musical genre classification of audio signals. *Speech and Audio Processing, IEEE transactions on* 2002; **10**(5):293–302.
3. Bogdanov, D., Serrà, J., Wack, N., Herrera, P., Serra, X.. Unifying low-level and high-level music similarity measures. *Multimedia, IEEE Transactions on* 2011;**13**(4):687–701.
4. Silla Jr, C.N., Koerich, A.L., Kaestner, C.A.. Feature selection in automatic music genre classification. In: *Multimedia, 2008. ISM 2008. Tenth IEEE International Symposium on*. IEEE; 2008, p. 39–44.
5. Lartillot, O., Toiviainen, P.. A matlab toolbox for musical feature extraction from audio. In: *International Conference on Digital Audio Effects*. 2007, p. 237–244.
6. Eronen, A., Klapuri, A.. Musical instrument recognition using cepstral coefficients and temporal features. In: *Acoustics, Speech, and Signal Processing, 2000. ICASSP'00. Proceedings. 2000 IEEE International Conference on*; vol. 2. IEEE; 2000, p. II753–II756.
7. Hall, M.A., Smith, L.A.. Practical feature subset selection for machine learning 1998;.
8. Kohavi, R., John, G.H.. Wrappers for feature subset selection. *Artificial intelligence* 1997;**97**(1):273–324.
9. Ghosh, A.K.. On optimum choice of k in nearest neighbor classification. *Computational Statistics & Data Analysis* 2006;**50**(11):3113–3123.
10. Devéze, B., Fouquin, M.. Datamining c4. 5–dbscan. 2005.
11. Sindhwani, V., Rakshit, S., Deodhare, D., Erdogmus, D., Principe, J.C., Niyogi, P.. Feature selection in mlps and svms based on maximum output information. *Neural Networks, IEEE Transactions on* 2004;**15**(4):937–948.
12. Keerthi, S.S., Shevade, S.K., Bhattacharyya, C., Murthy, K.R.K.. Improvements to platt's smo algorithm for svm classifier design. *Neural Computation* 2001;**13**(3):637–649.