Comparison of ML classifiers for Raga recognition

Hiteshwari Sharma, Rasmeet S.Bali

Department of Computer Science and Engineering, Chandigarh University, Mohali, Chandigarh

Abstract- Indian classical music is categorized into two major forms: Hindustani music and Carnatic music which are performed and practiced in North and South India. Raga is the fundamental concept in Indian music on which whole melody of a performance is based and comprises of group of swaras (musical notes) ranging from 5-8 notes. The present work focuses on Hindustani raga identification. We perform raga recognition on a set of ragas and collected dataset of live performances of these ragas, both vocal and instrumental, and performed recognition using machine learning classifiers. A comparative analysis is done between these classifiers based upon their parameters like average accuracy, precision, recall, kappa statistic and results are being formulated. The result shows that K-star algorithm performs better with raga identification with an accuracy of 93.38%.

Index Terms- Raga Recognition; Machine learning; Classifiers; Precision; Recall; Kappa statistic;K-star

I. Introduction

Hindustani music is oldest form of music and have no predetermined beginning or end but flowing without interruption through the fingers of the composer-performer**:** the tuning of the instrument merges imperceptibly with the elaboration of the melody, which may spin itself out for two, three or more unbroken hours[1]. It achieves a complexity both melodically and rhythmically. Moreover, movements in Indian classical music are on a one-note-at-a-time basis. Indian classical music is defined by two basic elements **-** it must follow a Raga (classical mode), and a specific rhythm, the Taal. In Indian classical composition, the music is based on drone i.e. a continual pitch that sounds throughout the concert, which is tonic. This acts as a point of reference for everything that follows a home base that the musician returns after a flight of improvisation. The result is a melodic structure that is easily recognizable, yet infinitely variable i.e. Raga [2].The Indian *raga* chooses between five to eight swaras (musical notes) and it forms the musical foundation on which an musical performance is based. Classical Indian music is characterized by seven main musical notes called the **'saptak'** viz. Shadja, Rishab, Gandhar, Madhyam, Pancham, Dhaivat and Nishad along with five intermediate notes known as altered notes or 'vikrit swaras'. Further the swaras can be played in three octaves, the first or lower octave starting from 130Hz; then middle octave starting from 260Hz; and the upper octave from 520 Hz. Moreover, Indian classical music allows the artist to improvise over the definitions of a raga to create their own personal performance of the raga. Due to this, two performances of the same raga by different artists may sound strikingly different to the novice ears, though they still maintain the defining qualities of that raga [3]. Raga can be identified by various parameters i.e Choice of notes, Ascending and Descending sequences (arohana and avarohana pattern), nature of inflexion on different notes (gamaka/meend), characteristic phrases (pakad). They are described as below:

1. Choice of notes: They are basically the combination of notes (swaras) used from saptak. Each raga has different set of swaras that constitutes it. Vadi swara is the king swara on which maximum focus is given on a performance. It is also known as most frequently occurring swara in a particular raga. It is followed by Samvadi (next in importance), then Anuvadi. The swaras that are not allowed in a particular raga are known as Vivadi swaras (enemy notes).

2. Arohana-Avarohana: This refers to ascending (upward) and descending (downward) sequence of notes in a raga.

3. Pakad: It is a group of notes that acts as catchy phrase for raga identification It is sung by the artist repeatedly after a particular interval.

4. Gamakas: Gamakas are better known as ornamentations used in Hindustani music system. These are inflexions and rapid oscillatory movements taken across swaras.

The rest of the paper is organized as follows: Section II describes the relevant literature for identifying Hindustani classical raga. In section III, the proposed algorithm and different steps for raga recognition is described. Section IV and V provides results and conclusions respectively.

II. RELATED WORK

Raga is identified based upon its fundamental characteristics: Arohana-Avarohana, Pakad, Gamakas, Choice of swaras, Vadi, Time and Season etc. These characteristics help in analyzing and framing different computational techniques for raga identification.

Pandey et.al.[2] proposed a system, Tansen,for automatic identification of raga based on Hidden markov model enhanced with string matching algorithm.Rajshri Pendekar et.al.[3] identified the raga by segmentation of audio signal via spectral flux and thereby identified raga by using its pitch frequency. LM Chelpa[4] proposed a fuzzy set theory for generation of alap patterns and have therefore identified different level of abstractions for raga identification.Chordia et.al.[5] derived Pitch-class Distributions (PCD)along with the Pitch-class Dyad Distributions (PCDD) from Harmonic Pitch Class Profiles (HPCP) and used these distributions for classification by using SVMs. In [6] Sreedhar et.al. created a database of ragas and used the scale of the raga performance as the similarity metric. Within a scale, notes are matched with the existing sets of notes in the ragas in the database. The closest raga in the database is given as the output for the test raga. Gulati et.al.[7] proposed pitch extraction procedure for carnatic music. Shetty et. al. [8] has identified raga based upon arohana-avorahana pattern on different ragas using neural network technique. Pranay et.al.[9] used scale independent raga identification using a Random forest classifier on swara histograms and achieved state-of-the-art results. Raga recognition is being done without the knowledge of the scale of the performance.

Inspired by the use of swara based feature for raga recognition, we analysed how these features extracted from different ragas are classified by different machine learning classifiers. The classifiers used for present study are K-star, C4.5, Random Forest and Bayesian network.

III. WORK DONE

There are no tentative methods for raga identification; usually there are well defined procedures by which experts recognize raga from a musical composition. It normally depends upon whether the person is trained or untrained. The untrained people usually associate two tunes for similarity and make their observations by comparing them whereas trained musicians perform raga identification based upon its fundamental characteristics. Even though professional musical veterans cannot perform this task of raga recognition with 100% accuracy.

Machine learning is a branch of artificial intelligence which deals with modeling of system to make predictions and behavior based on particular data. It comprises a complete set of classifiers which perform classification, prediction and regression based on different parameter. In this work, recognition is being performed on the four ragas i.e. Des, Bhupali, Yaman and Todi which are based on variation of ten basic Thaats, also known as musical scale using machine learning classifiers.Raga Des is under Bilawal thaat, Bhupali and Yaman are under Kalyan thaat and Todi is under Todi thaat. These four ragas are chosen for present analysis as they are fundamental ragas as suggested by musicians. Table I shows four ragas with their corresponding vadi swara, which is the most frequent note occurring in a particular raga.

Table I: Ragas and their vadi swara

S.no.	Raga	Vadi
1.	Des	R
2.	Bhupali	G
3.	Yaman	G
4.	Todi	D

Following are the sequence of steps in which raga recognition is being done.

A. Collection of Audio performances

Different vocal and instrumental performances are collected in four different ragas mentioned above and preprocessed for further analysis. Fifty live raga performances are taken for the present analysis.

B. Feature Extraction

The audio performances are converted into .wav (wave amplifier) extension and chroma features are extracted from them using MIR(Music Information Retrieval) toolbox in Matlab. A hop factor of 0.025 second is selected which gives 4719 frames. MIR is an open source toolbox which offers an integrated set of functions written in Matlab, dedicated to the extraction of musical features from audio files such as tonality, rhythm and other features like: Spectral Analysis, Chromagram analysis, Brightness, Centroid curve, Tempo, Pitch, Pulse etc.[12] Chromagram feature is being selected for the proposed scheme.

Chromagram is a visual representation of energies in the 12 semitones (or chromas) of the western musical octave namely C, C#, D, D#, E, F, F#, G, G#, A, A# and B. So, it basically depicts the distribution of energy in the twelve pitch classes From the chromagram, we extracted the semitone exhibiting maximum energy by using max function in MIR for each frame and get semitone sequences corresponding to that raga. Though these sequences of semitones might contain some identifying information about the raga, using them for raga identification is not sufficient since ragas are described by pattern of notes which might get overlapped with different western semitones which depends upon the musical scale selected by the musician. In our approach, we assume that we do not have information about scale selected for a particular raga performance. We must therefore find the mapping from the absolute frequency scale used by the chromagram to the corresponding scale of the musical piece, so that we can transcript the swara sequence. We will be taking the vadi (king swara) to convert our semitone sequence to a swara sequence [9].

C. Analysis of extracted features

Vadi swara, also known as king swara is selected for raga identification. Its magnitude and pitch frequency is relatively greater than notes (swaras).We had calculated magnitude of different notes in each frame of a chromagram. The note having the maximum magnitude in western scale is aligned with the vadi of that raga. One minute snippet of each raga is selected and the tabulated results are then analysed for further processing.

D. Classification using ML classifiers

The results tabulated in MS Excel sheet is imported in WEKA tool as .csv file. The WEKA tool provides a comprehensive collection of machine learning algorithms and data preprocessing tools to researchers. It allows comparisons between different machine learning methods on new data sets. Its modular, extensible architecture allows sophisticated data mining processes to be built up from the wide collection of base learning algorithms and tools provided. The workbench includes algorithms for regression, classification, clustering, association rule mining and attribute selection. WEKA has several graphical user interfaces like Explorer, Experimenter, Knowledge flow and simple Command Line Interpreter [12].

The data set of different ragas is classified using machine learning classifiers in WEKA. Classifiers used are Random forest classifier, C4.5, Bayesian network and K-star.They are described as below:

C4.5 classifier

C4.5 classifier builds decision trees from a set of training data using the concept of information entropy. Each training sample s_i consists of a p-dimensional vector $x_{1,i}$, $x_{2,i,.....,}x_{p,i}$

where x_j represent attributes or features of the sample, as well as the class in which s_i falls. At each node of the tree, C4.5 chooses the attribute of the data that most adequately splits its set of samples into subsets enriched in one class or the other.

Bayesian classifier

A Bayesian network is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed acyclic graph (DAG).Each node is associated with a probability function that takes as input a particular set of values for the node's parent variables and gives the probability of the variable represented by the node.

Random Forest classifier

Random forests (RF) are a combination of tree predictors and uses random selection of features to split each node growing an ensemble of trees and letting them vote for the most popular class. To grow these ensembles, often random vectors are generated that govern the growth of each tree in the ensemble. After a large number of trees is generated, they vote for the most popular class.

K-star classifier

K-star or K* is an instance-based classifier. The class of a test instance is based on the training instances similar to it, as determined by some similarity function. The K-star algorithm uses entropic measure, based on probability of transforming an instance into another by randomly choosing between all possible transformations.

In the figures, we have taken raga Des as a sample raga. Figures 1, 2,3 and 4 shown below comprises of spectrogram, Frame decomposition,chromagram and chroma magnitude values in matrix form of raga Des respectively.MIR toolbox functions are applied on raga performances which give the following results as shown in figures.

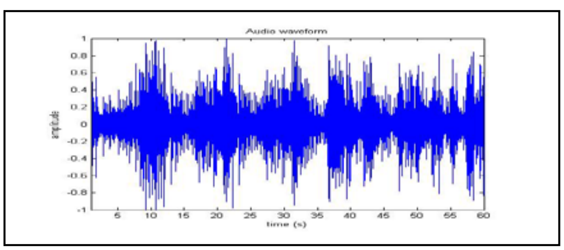

Fig. 1. Spectrogram of Raga Des

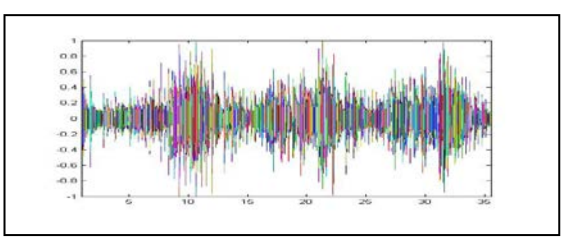

Fig. 2. Frame decomposition of Raga Des

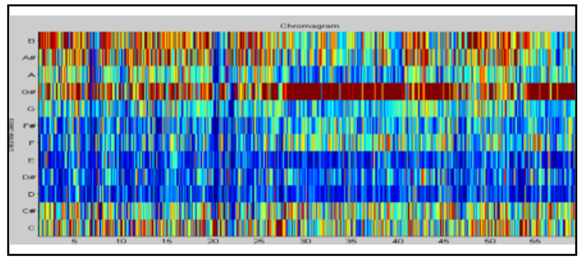

Fig.3. Chromagram of Raga Des

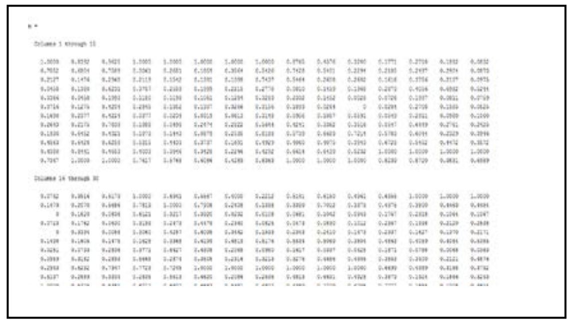

Fig.4.Chroma magnitude values of each frame

Fig.4.Chroma values of different frames of Raga Des

After performing chromagram analysis, magnitude and number of occurrences of different notes were found. We then calculated normalized frequency values by dividing the note values by total number of frames.

The above steps are performed on four ragas and results are tabulated. Fifty vocal and instrumental performances of each raga is taken for analysis purpose.

IV. RESULTS

After applying machine learning classifiers on the dataset of ragas, we performed analysis on its different parameters i.e. correctly and incorrectly classified instances, kappa statistics, mean absolute error etc. and accuracy achieved by each classifier. A comparative analysis was done among these classifiers based upon following parameter values. A ten fold cross validation is performed for testing, which means that on every tenth instance of the data set, testing is being performed. Table II and III shows comparative analysis among four machine learning classifiers with their corresponding parameter values.

Table II: ML Classifiers and their analysed parameter values for various ragas

S.no	Classifier	Average accuracy	Kappa Statistic	Mean absolute error
1.	C4.5	89.70	0.863	0.039
2.	Bayesian	86.76	0.826	0.054

3.	Random forest	92.64	0.90	0.051
4.	K star	93.38	0.912	0.023

The average accuracy is defined as the percentage of examples correctly classified over the total number of examples ,Kappa statistic measure the agreement of prediction with true class, True Positive (TP) rate is the proportion of examples which were classified as class x, among all examples which truly have class x, i.e. how much part of the class was captured. It is equivalent to Recall, False Positive (FP) rate is the proportion of examples which were classified as class x, but belong to a different class, among all examples which are not of class x, Precision is the proportion of the examples which truly have class x among all those which were classified as class x. The Frequency measure or F- measure is calculated by the formula: 2*Precision*Recall/ (Precision+Recall), which is a combined measure for precision and recall.

Table III: Detailed Accuracy of Classifiers by Class

S.no	Classifier	TP	FP	Precision	Fm
1.	C4.5	0.897	0.032	0.876	0.886
2.	Bayesian	0.868	0.036	0.861	0.851
3	Random Forest	0.91	0.029	0.888	0.876
4.	K star	0.934	0.017	0.934	0.934

After performing comparison of classifiers on ragas, it was observed that K star gives largest average accuracy of 93.38% on dataset of ragas, followed by random forest with 92.64%, then C4.5 with 89.70% and then Bayesian classifier with accuracy of 86.76%.Kappa statistic value in K * is 0.90, and hence denotes near to perfect classification. It was also observed that random forest classifier gives better accuracy as compared to K star when instances were less. But when the number of instances gets increased, K star performs better than random forest. However there is only a marginal difference between their accuracies. K star is an instance based learning classifier which uses entropy as a distance measure for classification problems. The distance function which determines how similar two instances are, and the classification function which specifies how instance similarities yield a final classification for the new instance. The computation of distance between two instances is motivated by information theory, according to which the distance between instances is defined as the complexity of transforming one instance into another.

The calculation of the complexity is done in two steps. First a finite set of transformations which map instances to instances is defined. A program which transform one instance (a) to another (b) is a finite sequence of transformations starting at a and terminating at b. The K star function can be calculated as:

$$K^* (y_{i,\,x}) = -\ln P^*(y_i, x) \qquad (1)$$

where data instances, *x*, are assigned to the class that occurs most frequently amongst the k-nearest data points, y_j, where j = 1,2…k. P* is the probability of all transformational paths from instance *x* to *y* [13].K-star therefore performs raga recognition by mapping two instances of raga and measures their distance, which is the complexity of transforming one instance into another.

Since Hindustani musical ragas follows set of rules, hence similar raga compositions will be transformed easily as compared to different ones. But again, raga performances are highly improvised and complex, due to which distances calculated by classifier gets bigger even in case of similar instances which leads to misclassification errors.

V. CONCLUSION AND FUTURE WORK

In this paper, we empirically evaluated and compared four machine learning classifiers on normalized dataset of ragas. Real time recognition of raga has been done. We achieved highest accuracy of 93.38% in case of K star learning classifier. The whole raga structure is very complex and despite having well defined rules, a musician, whether vocal or instrumental performer never follows these rules exactly. Moreover, Hindustani music is highly improvised as performer enjoys full freedom for any movements in raga which leads to misclassification errors.Fluctuations in human voice while live performances are taken into consideration and this has affected the accuracy of our work. We achieved a misclassification error of 6.61% and this can again be attributed to the fact that actual musical performances are improvised and do not follow the raga rules strictly. Hence, there is a lot of scope for improvement in results.

Future work lies in proposing a well defined and robust machine learning technique for raga recognition which gives maximum accuracy as compared to other classifiers and performs identification in same manner as done by humans followed by expansion of our dataset to many more ragas.

REFERENCES

[1] O.Kholkar,"North Indian Classical Music", 2001
[2] G. Pandey, C. Mishra, and P. Ipe, "Tansen: A system for automatic raga identification" Proc. of Indian International Conference on Artificial Intelligence,2003, pp. 1350-1363.
[3] R.Pendekar,S.PMahajan and R.Mujumdar,,"Harmonim Raga Recognition,"International journal of Machine learning and Computing,vol.3,No.4, August 2013
[4] L.M.Chelpa,"A fuzzy set theoretic framework towards computer generation of alap in Hindustani music", NITR, Mumbai.
[5] P.Chordia and A.Rae,"Raag recognition using Pitch-class and pitch-class dyad distributions",Proc. of ISMIR,2007,pp. 431-436.
[6] R.Sridhar and T.Geetha,"Raga identification of carnatic music for music information retrieval",International Journal of Recent Trends in Engineering, vol.1,no.1,2009,pp.571-574
[7] S.Gulati and P.Rao,"A Survey of raag recognition techniques and improvement to the state of the art",Sound and Music Computing,2011
[8] S.Shetty and K.Achary,"Raga mining of Indian music by extracting arohana-avarohana pattern",International Journal of Recent trends in Engineering,vol.1,no.1,2009
[9] P.Dighe and H.karnick,"Swara histogram based structural analysis and identification of Indian classical music", ISMIR,2012

[10] H.Sahasrabuddhe,"Can mindware and software meet",Indian International Centre,1994,New Delhi

[11] Ram Avtar Vir, "Theory of Indian music",Pankaj Publications, New Delhi,1999

[12] O.Lartillot and P.Toiviainen, "A Matlab Toolbox for Musical Feature Extraction From Audio",International Conference on Digital Audio Effects, Bordeaux, 2007.

[13] R.Bouckaert and E.Frank,"Weka Manual for version 3-7-8", University of Waikato, New Zealand

[14] J.E Cleary and L.E Trigg,"K STAR: An Instance-based Learner Using an Entropic Distance Measure",University of Waikato, New Zealand

[15] F.Diniz and F.Milton,"A Facial Recognition System Based on Techniques of Principal Component Analysis and Autofaces with K-NN, K-Star and Random Forest Classifiers", Federal University Rural of Semiarid,2013

[16] H.Sharma and R.Bali," Raga Identification of Hindustani music using soft computing techniques",Proc. of RAECS, IEEE conference,Chandigarh,April 2014.

AUTHORS

First Author – Hiteshwari Sharma, Department of Computer Science and Engineering, Chandigarh University, Mohali, Chandigarh, hiteshwarisharma@gmail.com

Second Author – Rasmeet S.Bali, Department of Computer Science and Engineering, Chandigarh University, Mohali, Chandigarh, rasmeetsbali@gmail.com