Indian Classical Raga Identification using Machine Learning

Dipti Joshi, Dr. Jyoti Pareek, Pushkar Ambatkar

Department of Computer Science, Gujarat University, Ahmedabad, Gujarat, India

Abstract

Ragas demonstrate the pride of Indian classical music. Raga is the original musical form in Indian classical music. It consists of set of swaras (lyrical notes) that made up of various characteristics as a melodious conception which is played by the instruments and the singer. Based on the features of the raga, the Indian classical music is separated into two Parts: Hindustani (North Indian) and Carnatic (South Indian) classical music. Our experiment is concentrate on Hindustani classical music. In our experiment, K Nearest Neighbor (KNN) and Support vector machine (SVM) classifiers are used on the raga dataset of Yaman and Bhairavi to achieve classification and identification of the raga. We have done accurate outcomes with both KNN and SVM classifiers.

Keywords

Raga identification, Feature extraction, Machine Learning, KNN, SVM

1. Introduction

Indian classical music is the music of the Indian subcontinent. Raga or Raag hold a prominent position in Indian Classical Music. A Raag is a collection of musical notes that, when sung or performed on a musical instrument, are quite attractive. Raga recognition comprises of methods that define and classify notes from a piece of music into a suitable raga. In Hindustani classical music, Ragas is a very significant idea and express the moods and sentiments of concert. The classification of ragas comes only after an enough amount of exposure as it is an intellectual process.[1] Any of the attributes of ragas have to be translated into appropriate characteristics for automated recognition.

ISIC'21: International Semantic Intelligence Conference, February 25-27, 2021, New Delhi, India
joshidipti1408@gmail.com (D. Joshi)
0000-0001-9166-4555 (D. Joshi)

The characteristics of Ragas are based on Indian Classical Music techniques, which blend notes with the following features to qualify as a Raga.

Notes (swaras)

There have to be at least 5 or 7 notes (swaras) in a Raag. The primary seven notes are S (Sa), R (Re or Ri), G (Ga), M (Ma), P (Pa), D (Dha), N (Ni).

Aaroh and Avroh

Each Raga or Raag is composed of a "Aaroh" that implies swaras scale up and a "Avroh" that implies swaras scale down.

Vadi and Samvadi

Each raag consisting of "Vadi" means main notes and "Samvadi" means supporting swaras.

Gamakas

It has a constant frequency rate. Notes in a raga are a series of continuous (back and forth movement in a rhythm) variation, such sort of notes are known as Gamakas.

Pakad

A set of Swaras which are distinctively recognizes a raga. There is a particular Pakad for each raga.

Tala

Tala refers to a rhythmic form, which is constructed from variety of beats.

Thaat

Thaat is used in raga classification. There are unique ten Thaats namely Kalyan, Bilawal, Bhairav, Khamaj, Poorvi, Marwa, Kafi, Asawari, Bhairavi, and Todi.

The piece of music has to be converted to Swara for classification. Due to the following factors, there are several difficulties in converting the piece of music in Swara [1]

1. During any performance, a music part is made up of many instruments.
2. The notes in Indian classical music are on a relative scale.
3. In a raga, there is no static initial Swara.
4. In Indian music, the notes do not have a predetermined frequency rate.
5. In classical music, the series of the swaras in the particular ragas is not static as it allows various innovations.

The key purpose of raga recognition is that it will provide a good start for Hindustani music information retrieval and it allows us to predict the raga's performance and accuracy. Besides this, for music analysis, we can also create a playlist focused on ragas.

2. Related work

In this section, we have reviewed different work done by other authors and analyzed their work for future scope. We have tried to give an analysis of various classifiers, their relevance and performance for Raga identification. In [2] Sharma, Hiteshwari, Bali, Rasmeet S, Raga identification have been done on the four ragas like Des, Bhupali, Yaman, and Todi - dataset of live performances of both voice-based and instrumental, and executed identification using pitch class profile and n-gram histogram machine learning classifiers. For the pitch class profile, they received 83.39% accuracy and 97.3% for the n-gram histogram. In this paper [3] Ekta Patel and Savita Chauhan, have used the MATLAB toolbox for extracting track functions. A machine learning tool WEKA is used which works on .arff file format. Bayesian net, Naive Bayes, Support vector machine (SVM), J48, Decision table, Random forest classifiers on Bhairav, Yaman, Shanakara, Saarang dataset. The predominant demanding situations are the complicated variables like pitch and mood in the music track, skipping greater tones, the transformation of various dataset parameters and Raag. The effects are as compared before and after discretization, though in this Raag music identification, the accuracy of the possibility-based classifier is greater. It shows that a probability-based classifier gave accurate results. Comparatively, Bayesian Net provides better performance. [4] Hiteshwari Sharma and R. S. Bali have recognized various key variables for raga classification and Soft computing fuzzy sets technique for recognition of raga. They used a dataset of five ragas like Des, Bhupali, Yaman, Todi, and Pahadi with three parameters as time, dirgaswaras, and vadi and they have achieved reasonable accuracy as well. [5] G. Pandey, C. Mishra, and Paul Ipe introduced the Hidden Markov Model and pakad matching on the dataset of two ragas Bhupali and Yaman kalian. They have achieved an 77% accuracy with basic HMM and 87% accuracy with both HMM and Pakad matching methods. In [6] Muhammad Asim Ali and Zain Ahmed Siddiqui, their research was based on Automatic Music Genres Classification using Machine Learning. For which they have used algorithms like the K Nearest Neighbor (KNN) and Support Vector Machine (SVM) to anticipate the genre of songs. Using the GTZAN dataset, which has a wide range of ten genres, such as blues, hip-hop, jazz, classical, metal, reggae, country, pop, disco, and rock, they gathered musical data. They used the data set of 1000 songs. The above comparison shows that SVM is a more efficient classifier than KNN. [7] SnigdhaChillara, Kavitha A, Shwetha A Neginhal, Shreya Haldia, Vidyullatha K, proposed to solve the classification problem and comparison among some other models using the Free Music Archive small (fma_small) dataset. In that, two sorts of inputs were given to the models. Wherein CNN models used the spectrogram images and .csv file for Logistic Regression and ANN model used audio features stored in. They have received 88.5% accuracy using CNN on the spectrogram based model which is quite good compared to different algorithms used by other authors.

3. Work Done

In this section, we will discuss different characteristics of audio and Machine Learning algorithms like K Nearest Neighbor and Support Vector Machine in a brief way.

3.1 Feature Extraction

Each audio signal comprises of several features. But, it requires fetching the characteristics that are suitable for the issue that we want to solve. The method of fetching characteristics to apply for the study is referred to as feature extraction. We will have a brief idea about some of the characteristics below, in detail.

Power Spectrogram

A spectrogram is a graphical demonstration of the spectrum of frequencies of a signal as it differs with time. When it is used with an audio signal, spectrograms are sometimes referred to as the sonographs, voiceprints or voicegrams. To determine the raga, we are using the mean of spectrogram to get which tone/ pitch is used more.

MFCC- Mel-Frequency Cepstral Coefficients

This feature is one of the most necessary techniques to extract attributes of an audio signal and it is used mostly when we are working on audio signals. The mfccs of a signal are a set of characteristics (approximately (10–20)) which in brief illustrates the general form of a spectral cover.

Spectral Centroid

It shows that the "center of mass" that considered the weighted mean of the frequencies present within the sound. If it gets the equal frequencies in tune for a particular time span, then the spectral centroid might be around the center and if there are excessive frequencies at the end of the sound then the centroid tends to be closer to its end.

Zero-Crossing Rate

The rate at which sign varies is known as the zero-crossing rate. Zero crossing rate is the rate wherein the signal varies from positive to negative and vice versa. Speech recognition and music information retrieval are being commonly used in Zero crossing rates. It has excessive values for loud and noisy sounds like in metal and rock.

Roll-Off Frequency

Particularly Roll-off suggests the activity of a particular sort of channel; one planned to Roll-off frequencies raised or lowered at a certain point. It is called roll-off as the method is progressive.

Spectral Bandwidth

A radiated spectral quantity is not less than half its maximum value in spectral bandwidth. It determines the extent of the Spectrum.

This is an interval difference between lower and higher frequency.

3.2 Machine Learning Algorithms

During our analysis, we have found that the Supervised Machine Learning approach might be a good fit for our problem.

We have tried to implement various classification algorithms and found that K Nearest Neighbor and Support Vector Machine is quite appropriate for our experiment.

3.2.1 K Nearest Neighbor (KNN)

K Nearest Neighbor (KNN) is a supervised learning method. It is the simplest but robust algorithm that is applied for both regression and classification problems. To build a prediction, the KNN algorithm uses the whole dataset in which we attempt to classify data points to a particular category with the help of the training set.

3.2.2 Support Vector Machine (SVM)

Support vector machine (SVM) is also a supervised learning method that is usually used for classification. A hyperplane that clearly divides the sampling points with various labels is identified by this algorithm. It separates sample points of both labels and class on different sides of the hyperplane.

4. Dataset and Strategy

In this experiment, we have used the dataset of Audio files. The dataset was created by extracting 60-second audio clips from the internet. For music and audio analysis, a python package Librosa is used. It provides the segments that are required for creating music information retrieval systems. Another open-source machine learning library is Scikit-learn which supports both, supervised and unsupervised learning methods. It also provides a variety of tools for model selection, data pre-processing, model fitting, and estimation. In this experiment, we have chosen Yaman and Bhairavi Raga. We have split them into 60-sec frames which allow the computer to work only on the specific part of the song like Pakad, Aaroh, and Avroh and remove other noise and empty fields from the audio. Using Librosa Library, we have created a .CSV file to save all the features like Mfcc, Spectrogram, Bandwidth, Centroid, zero-crossing, and Roll-off which are to be extracted from the audio file. Below is the flow of the process.

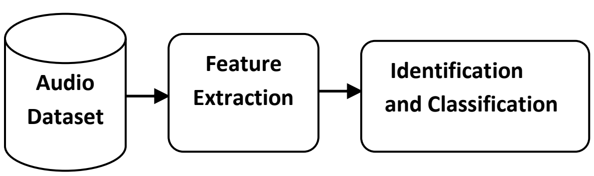

Figure 1: Process of Raga Identification [8]

This is the visual representation of the classification procedure in which features are taken out from the audio frame and compared with the weight of the closer mean. Using Scikit Learn library, we have implemented KNN and SVM algorithms on the .CSV data file. We found that both KNN and SVM fit best for the classification of raga other than Logistic Regression.

5. Result and Discussion

In our experiment, we have chosen Ragas like Yaman and Bhairavi. We have used vocal-instrument Dataset consisting of 341 audio clips, out of which, 194 audio clips are of Yaman and 147 audio clips are of Bhairavi.

The tables below show the accuracy of KNN and SVM algorithms.

Table.1 Results of Raga Identification accuracy Using KNN

KNN	Train/Test Ratio	Classification Accuracy
1	80 / 20	98%
	60 / 40	94%
	40 / 60	93%
2	80 / 20	97%
	60 / 40	93%
	40 / 60	92%
3	80 / 20	97%
	60 / 40	94%
	40 / 60	93%
4	80 / 20	95%
	60 / 40	93%
	40 / 60	93%

Table.2 Results of Raga Identification accuracy using SVM

Classifier	Train/Test Ratio	Classification Accuracy
SVM	80/20	95%
	60/40	95%
	40/60	94%

In both the tables given above, we have implemented two different classifiers i.e. KNN and SVM. Using KNN, we can observe that the accuracy of raga identification is varied for different Neighbour values. We have divided the

Train/ Test ratio into 3 different categories that are 80/20, 60/40, and 40/60 respectively. For all KNN values, we got the highest accuracy in Train/Test ratio 80/20 i.e. For KNN value 1 to 4; we received 98%, 97%, 97%, and 95% accuracy respectively. Using SVM, we got 95% accuracy for the 80/20 Train/Test ratio.

From this comparison, we can conclude that KNN with all Neighbor values for Train/Test ratio 80/20 gives the highest accuracy which is better in comparison with SVM.

6. Conclusion and Future work

A short introduction on raga and its attributes are considered. Prior policies for raga classification and recognition are observed with their data records, implementation applications, correctness, and problems. In this paper, we have discussed the classification of different ragas like Yaman and Bhairavi by applying K-Nearest-Neighbor (KNN), Support vector machine (SVM) machine learning algorithms. We have obtained good results with KNN and SVM, but in our experiment, KNN seems to be performing slightly better.

In the future, we will expand our dataset with many other Ragas for acquiring more accurate results and also implement other classifiers for detecting ragas in a well-defined manner.

References

[1] Vijay Kumar, Harit Pandya, C.V. Jawahar, “Identifying Ragas in Indian Music”, 22nd International Conference on Pattern Recognition 2014.

[2] H. Sharma and R. S. Bali, “Comparison of ML classifiers for Raga recognition,” Int. J. Sci. Res. Publ., vol. 5, no. 10, pp. 1–5, 2015.

[3] E. Patel and S. Chauhan, “Raag detection in music using supervised machine learning approach,” Int. J. Adv. Technol. Eng. Explor., vol. 4, no. 29, pp. 58–67, 2017.

[4] H. Sharma and R. S. Bali, “Raga identification of Hindustani music using soft computing techniques,” 2014 Recent Adv. Eng. Comput. Sci. RAECS 2014, pp. 6–8, 2014.

[5] G. Pandey, C. Mishra, and P. Ipe, “Tansen: A system for automatic raga identification,” Indian Int. Conf. Artif. Intell., pp. 1350–1363, 2003.

[6] Muhammad Asim Ali and Zain Ahmed Siddiqui ,”Automatic Music Genres Classification using Machine Learning”, International Journal of Advanced Computer Science and Applications, Vol. 8, No. 8, 2017.

[7] S. Chillara, A. S. Kavitha, S. A. Neginhal, S. Haldia, and K. S. Vidyullatha, “Music Genre Classification using Machine Learning Algorithms : A comparison,” no. May, pp. 851–858, 2019.

[8] Kalyani C. Waghmare and Balwant A. Sonkamble, “Raga Identification Techniques for Classifying Indian Classical Music: A Survey”, International Journal of Signal Processing Systems Vol. 5, No. 4, December 2017.