Classifying Indian Classical Music(Carnatic) and Western Classical Music

Author: Karishma Kamalahasan

ABSTRACT

Indian classical music is the music of the Indian subcontinent and is one of the oldest kinds of music. Although Indian music has progressed significantly, the fundamental components appear to remain the same as they were two thousand years ago. The raga is at the center of Indian classical music, and it is the language of the spirit. Novel Indian music, on the other hand, has been heavily impacted by Western music notation, which is based on equal mean tones, temperament scales, and the universal priority of harmony. Indian and Western music differs not just in terms of culture, but also in terms of fundamental structures, scales, and tuning. New musical innovations are burgeoning in the current decade with an incredible number of fusions of Western Classical music and Indian Classical music. While fusions are a way of celebrating two different art forms, is the originality of the classical forms fading away? This project aims at studying the classification of Indian Classical Music and Western Classical music. A simple feedforward neural network was created to classify the audio data. The audio classification is not limited to the style of music. The model also classifies the audios according to the instruments and the different ensembles at play. A testing accuracy of 75.00 percent and a training accuracy of 97.9 percent was achieved through the model. This project is an attempt to encourage musical fusions and to help create awareness among the community about the ancient and widely practiced art of Carnatic Music.

I. INTRODUCTION

Music is an art devoid of language barriers. In olden days, music has been an essential cultural and social factor throughout human history. Initially, music has been looked at from a single cultural perspective but with passing times, music is seen from a musical perspective of different cultures paving the path to multiculturalism [1].

Indian classical music, originated in South Asia is a rich tradition found in all corners of the world. Its origin dates back to sacred Vedic scriptures over 6,000 years ago where chants developed a system of musical notes and rhythmic cycles. This music is very closely connected to nature, taking inspiration from natural phenomena including the seasons and times of the day to create 'ragas' or musical moods and plenty of time cycles or 'taals' that are further codified [2].

Carnatic music derives its name from 'Karnātaka Sangītam', a Sanskrit term, which denotes "traditional" or "codified" music [3]. Carnatic music, the south Indian classical music, has a rich history. It was developed in the south Indian states of Tamil Nadu, Kerala, Andhra Pradesh, and Karnataka which are known for their strong presentation of Dravidian culture. Purandardas (1480-1564) is the father of Carnatic music as he devised the method of codification of Carnatic music [4]. Venkat Mukhi Swami is regarded as the grand theorist of Carnatic music ,who also developed "Melankara", the system for classifying south Indian ragas [5].

Western music is a music of the West which celebrates the culture of Americans, Europeans, and other societies established by European immigrants. The history of Western music date back to the frontier era of the 19th century. It uses major and minor scales and equal temperament notes, unlike Indian Classical music Western scales usually consist of seven notes and five variations that are arranged in an order of increasing pitch to form a scale or a gamut, referred to as an "octave" in the Western music [6].

II. DATASET

In this section, the fundamental part of the experiment is discussed: Data collection and usage.

Due to the unavailability of a standard dataset consisting of Carnatic classical music and Western Classical music along with the ensembles and instruments used to play the music, a custom dataset was developed.

To test and train the model on Carnatic music, data was collected from Saraga's research dataset of Indian Art Music(1.5-Carnatic) [7]. From the above exhaustive dataset, nine songs were collected randomly from each of the four categories: violin, mridangam(left), ghatam and vocals. Each song originally ranged from 30 seconds to 50 minutes, which was shortened to 10-15 seconds each. The above task was done in a randomized fashion using iMovie and Google's random number generator (RNG). Using the RNG, the range (in seconds) was entered, and 10-15 seconds of audio was used from that generated number. This method of extracting audio recordings of 10-15 seconds each creates uniformity and avoids inherent biases.

Similarly, to assemble the Western Classical music, data was collected from MusicNet (1.0) [8]. Instead of utilizing the entire dataset, six different ensembles of western classical music were organized: Solo Piano, String Quartet, Solo Violin, Wind Quintet, Solo Cello, and Accompanied Violin. To organize the data in the same style as the Indian music, nine songs were selected randomly from each of the six categories and extracted 10-15 seconds of audio randomly using Google's automated number generator.

All the audio files were in .wav format with an average sample rate of 24 bits per second or 48 kHz.

After having installed and having organized the dataset, a .csv file was designed with all the relevant information about each music piece including file name, salience, folder number, classID and class. The laying out of the dataset in this manner helps the classification model work more efficiently as it can read the file information using pandas directly, and it also helps the programmer reduce the lines of code in terms of the pathways for every file.

III. CODE AND ML MODEL

The code is a neural network with a Keras and TensorFlow backend [9]. It begins with importing all the necessary libraries including NumPy, pandas, librosa, matplotlib and Keras modules and functions for model building. Then dataset is loaded and read the .csv file using pandas. After which a single audio file is extracted from each class to display its waveform.

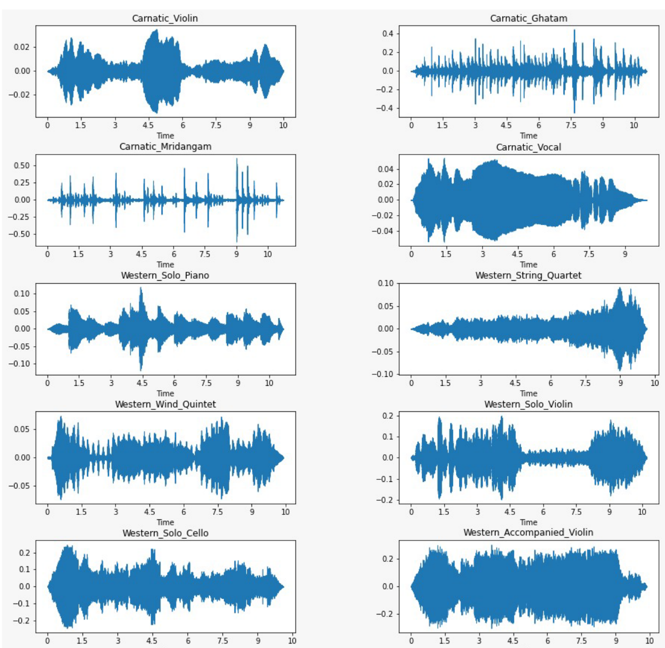

Figure 1: *Waveforms of 10 audio files from each class The waveforms are graphical representations with the amplitude as the Y axis and time as the X axis. [from the 1st column downwards: Carnatic_Violin, Carnatic_Mridangam, Western_Solo_Piano, Western_Wind _Quintet, Western_Solo_Cello][from the 2nd column downwards: Carnatic_Ghatam, Carnatic_Vocal, Western_String _Quartet, Western_Solo_Violin, Western_Accompanied_Violin]*

Soon after which features were extracted from the data using librosa and 'scipy.io.'. The 'librosa.load' function is used to convert the sampling rate to 22.05 kHz ,and also converts the audio signal from stereo to mono.

The audio waveform was also extracted in the MFCCs format (Mel Frequency Cepstral Coefficients) using Librosa's MFCC function. MFCCs of an audio signal comprise of certain (10-20) features which describe the overall shape of a spectral envelope.

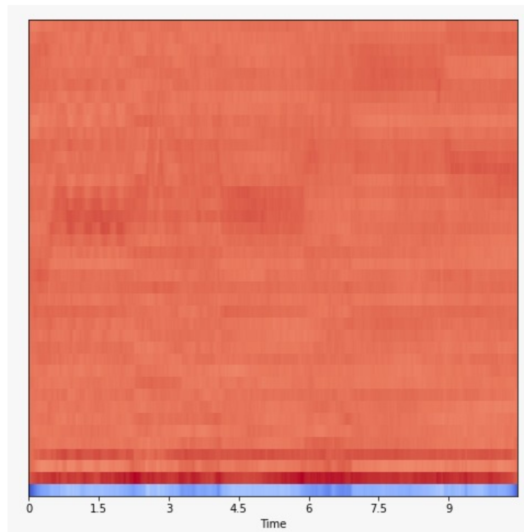

Figure 2: *MFCC of first audio file in fold . X axis: Time(seconds), Y axis: MFCC's(coefficients in numbers)*

After extracting the MFCCs for all the audio files, the features were displayed in the form of a 40-element array to depict the 40 MFCCs.

```
array([-5.41517151e+02,  1.03288857e+02, -2.09685421e+01,  4.51897926e+01,
       -7.31762743e+00, -1.06241798e+01, -1.04500074e+01, -9.71995735e+00,
       -3.11487412e+00, -1.09566174e+01, -1.54824762e+01, -4.04470110e+00,
       -7.72584963e+00, -5.83546305e+00, -1.89279795e+00,  4.21685457e+00,
       -7.58998919e+00,  5.25713253e+00, -7.58154809e-01, -1.27665873e+01,
       -2.89532959e-01, -1.49172628e+00,  1.10488358e+01,  1.37288427e+01,
        1.06224051e+01,  1.29149094e+01,  3.51227188e+00,  8.99843121e+00,
        1.01063786e+01,  9.53975391e+00,  4.43502776e-02, -1.01534195e+01,
       -7.75432301e+00, -3.96182323e+00,  7.05239153e+00,  3.21995211e+00,
       -2.52421021e-01, -1.84600562e-01, -1.97548246e+00, -5.05416965e+00],
      dtype=float32)
```

Figure 3: *40 element array of MFCCs*

Prior to model building, the data set was split into training and testing randomly using 'train_test_split' function, which split the 90 audio files into 50-training and 40-testing. Post splitting the dataset, a simple feedforward neural network architecture was built in a sequential manner, and it falls under the Multi-layer Perceptron (MLP) models. An MLP is a network that maps nonlinear inputs and outputs. The model was built using Dense, Dropout and Activation keras layers.

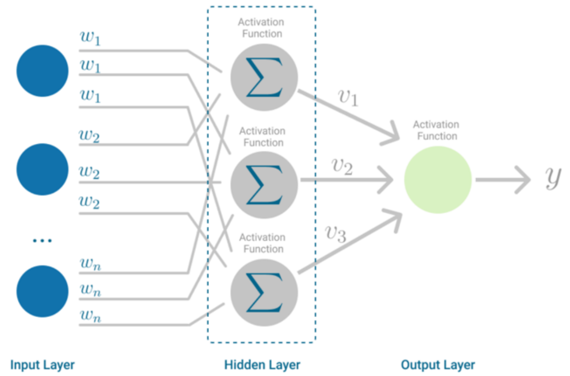

Figure 4: *MLP Model: Works with Backpropagation to minimize cost function.[10]*

```
Model: "sequential_2"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense_4 (Dense)              (None, 256)               10496
_________________________________________________________________
activation_4 (Activation)    (None, 256)               0
_________________________________________________________________
dropout_3 (Dropout)          (None, 256)               0
_________________________________________________________________
dense_5 (Dense)              (None, 256)               65792
_________________________________________________________________
activation_5 (Activation)    (None, 256)               0
_________________________________________________________________
dropout_4 (Dropout)          (None, 256)               0
_________________________________________________________________
dense_6 (Dense)              (None, 10)                2570
_________________________________________________________________
activation_6 (Activation)    (None, 10)                0
=================================================================
Total params: 78,858
Trainable params: 78,858
Non-trainable params: 0
_________________________________________________________________
```

Figure 5: *Model Summary*

The pretraining accuracy was 15.0000%. After which the data was trained with 134 epochs which led to a training accuracy of 97.90% and a testing accuracy of 75.00%.

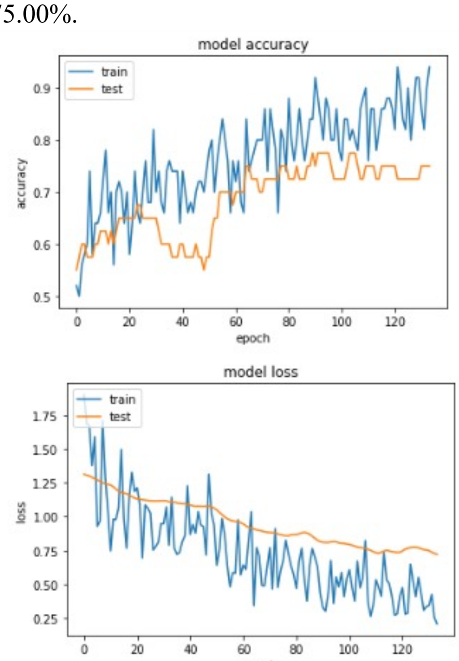

Figure 6: *Model Accuracy and Model Loss*

IV. ANALYSIS THROUGH COMET.ML

To get a more in-depth analysis of the model and its accuracy versus loss, Comet.ml [11] was imported and all the information produced was logged during the compilation and execution of the code. The detailed analysis is given in the following graphs.

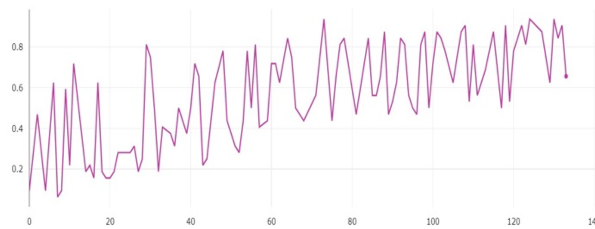

Figure 7: *Training Accuracy. X axis: epochs, Y axis: accuracy(%)*

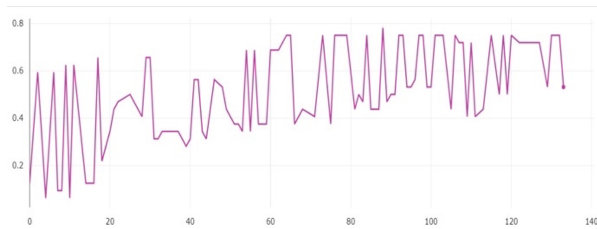

Figure 8: *Testing Accuracy. X axis: epochs, Y axis: accuracy(%)*

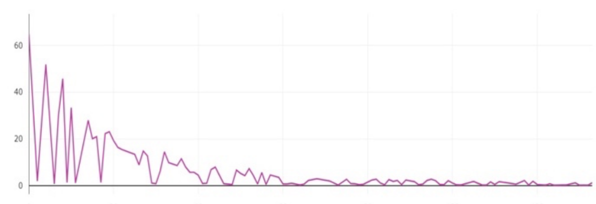

Figure 9: *Training Loss. X axis: epochs, Y axis: loss(%)*

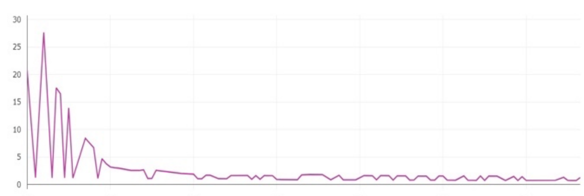

Figure 10: *Testing Loss. X axis: epochs, Y axis: loss(%)*

V. FUTURE WORKS

This project is still at its infant stage in terms of achieving its end goal. Furthermore, this work can be expanded in various ways: identifying if the authenticity of original classical music is retained in emerging genres and fusion music; comparative analysis of fusion music, generating mathematical combinations of various genres of music; identifying the different genres in fusion music.

V. ACKNOWLEDGEMENTS

I would like to acknowledge Polygence for giving me the opportunity to explore a project in my interested career path. I am especially very thankful to my mentor, Mohamed Eltaeb, for guiding me through my research. I would like to express my deep gratitude to Dr. Ramaprasad Poojary for his valuable and constructive suggestions during the planning and development of this research work.

VI. REFERENCES

[1] *The Cultural Connection*. UUA.org. (2022). https://www.uua.org/re/tapestry/resources/music/chapter8/129383.shtml.

[2] *An introduction to: Indian classical music | Making Music*. Makingmusic.org.uk. (2022). https://www.makingmusic.org.uk/resource/introduction-indian-classical-music#:~:text=What%20is%20Indian%20classical%20music%3F,musical%20notes%20and%20rhythmic%20cycles.

[3] *Carnatic Music - Carnatic Classical Music - Indian* . Cultural India https://www.culturalindia.net/indian-music/carnatic-music.html.

[4] *Indian Music - Genres, History & Evolution.* Cultural India. https://www.culturalindia.net/indian-music/

[5] *Carnatic Vocals Classes | Saanwee Performing Arts | United States*, https://www.saanwee.com/carnatic-vocals-for-kids-and-adults

[6] Khillar, S. (2022). *Difference Between Indian Music Notes and Western Music Notes | Difference Between*. Differencebetween.net http://www.differencebetween.net/miscellaneous/entertainment-miscellaneous/difference-between-indian-music-notes-and-western-music-notes/.

[7] Bozkurt, B., Srinivasamurthy, A., Gulati, S., & Serra, X. (2018). *Saraga: research datasets of Indian Art Music (1.5) [Data set].* Zenodo. https://doi.org/10.5281/zenodo.4301737

[8] John Thickstun, Zaid Harchaoui, & Sham M. Kakade. (2016). *MusicNet (1.0) [Data set]*. Zenodo. https://doi.org/10.5281/zenodo.5120004

[9] Mendels, Gideon. (November 18, 2019) *"How to Apply Machine Learning and Deep Learning Methods to Audio Analysis."* Medium. Towards Data Science, https://towardsdatascience.com/how-to-apply-machine-learning-and-deep-learning-methods-to-audio-analysis-615e286fcbbc.

[10] *Multilayer Perceptron Explained with a Real-Life Example and Python Code: Sentiment Analysis*. Medium. (2022). https://towardsdatascience.com/multilayer-perceptron-explained-with-a-real-life-example-and-python-code-sentiment-analysis-cb408ee93141.

[11] *Comet | Build better models faster*. Comet. (2022). https://www.comet.com/site/.