International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Indian Music Genre Classification

Major Project Report submitted in partial fulfillment of the requirements for the award of the Degree of B.E in Computer Science and Engineering

By

Komal Rajput
160618733150

Shagufta Ahmed
160618733169

T. S. Laasya
160618733174

Under the Guidance of

RAJASHEKAR SHASTRY
Asst.Prof of Dept of CSE
Department of Computer Science & Engineering

Department of Computer Science and Engineering Stanley College of Engineering & Technology for Women (Autonomous)

Chapel Road, Abids, Hyderabad – 500001

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

(Affiliated to Osmania University, Hyderabad, Approved by AICTE, Accredited by NBA & NAAC with A Grade)

2021-22

Stanley College of Engineering & Technology for Women

(Autonomous)

Chapel Road, Abids, Hyderabad – 500001

(Affiliated to Osmania University, Hyderabad, Approved by AICTE, Accredited by NBA & NAAC with A Grade)

CERTIFICATE

This is to certify that major project report entitled Indian Music Genre Classification being submitted by

Komal Rajput
160618733150
Shagufta Ahmed
160618733169
T.S. Laasya
160618733174

in partial fulfillment for the award of the Degree of Bachelor of Engineering in Computer Science & Engineering to the Osmania University, Hyderabad is a record of bonafide work carried out under my guidance and supervision. The results embodied in this project report have not been submitted to any other University or Institute for the award of any Degree or Diploma.

Guide **Head of the Department**

Rajshashekar Shasrty Dr Y V S S Pragathi

Asst. Prof of CSE Prof & HoD, Dept of

CSE

Project Coordinator **External**

Examiner

D Radhika
Asst Prof of CSE

Shuguftha Fatima
Asst Prof ofCSE

International Journal for Innovative
Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

DECLARATION

We hereby declare that major project work entitled ***Indian Music Genre Classification*** submitted to the Osmania University, Hyderabad, is a record of original work done by us. This project work is submitted in partial fulfilment of the requirements for the award of the degree of the B.E in Computer Science and Engineering.

Komal Rajput	160618733150
Shagufta Ahmed	160618733169
T.S. Laasya	160618733174

ACKNOWLEDGEMENT

First of all, we are grateful to **The Almighty God** for establishing us to complete this Major Project. We pay our respects and love to our parents and all our family members and friends for their love and encouragement throughout our career.

We wish to express our sincere thanks to Sri. Kodali Krishna Rao, Correspondent and Secretary Stanley College of Engineering & Technology for Women, for providing us with all the necessary facilities.

We place on record, our sincere gratitude to Prof. Satya Prasad Lanka, Principal, for his constant encouragement.

We deeply express our sincere thanks to our Head of the Department, Prof Y V S S Pragathi, for encouraging and allowing us to present the Major Project on the topic Indian music genre classification **at** our department premises for the partial fulfillment of the requirements leading to the award of the B.E. degree.

It is our privilege to express sincere regards to our project guide Rajshashekar Shasrty for the valuable inputs, able guidance, encouragement, whole-hearted co-operation and constructive criticism throughout the duration of our project.

We take this opportunity to thank all our faculty, who have directly or indirectly helped our project. Last but not least, we express our thanks to our friends for their co-operation and support.

ABSTRACT

In this project, we are going to build an application that classifies the genre of the music. The audio files will be given as input and the classified genre will be given as output. The INDIAN genre collection dataset was collected in 2000-2001. It consists of 600 audio files each having 30 seconds duration. There are 6 classes (6 music genres) each containing 100 audio tracks. Each track is in .wav format. It contains audio files of the following 6 genres:

- Carnatic
- Ghazal
- Semi-Classical
- Sufi
- Bolly-pop

There are various methods to perform classification on this dataset. Some of these

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

approaches are: Multiclass support vector machines, K-means clustering, K-nearest neighbor, Artificial neural network. The main domain used in this project is Machine Learning, and Deep learning using python. Summary: In this music genre classification project, we have developed a classifier on audio files to predict its genre. We work through this project on the INDIAN music genre classification dataset. This tutorial explains how to extract important features from audio files.

Keywords: sufi, classes, classification

Tables of Contents

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

List of Figures

CHAPTER 1

INTRODUCTION

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

1.1 About Project

Due to the increase in data consumption, music which has been the source of entertainment for centuries uncounted, has also take the online route to reach its patrons. The sheer volume of data being uploaded on these cloud platforms however, are not automatically assigned a genre. Automatic music genre classification is the essence for organizations that sell music or let users stream their content for a fee. The music database as a result would need to provide accessibility to its users to search, retrieve, and make automatic recommendations to the listeners based on their taste. Due to the very difficult task of extraction of acceptable audio features, music classification is regarded as a difficult undertaking. While unlabelled data is widely available, there is a scarcity of music with properly classified genre. The two phases in music genre categorization are the extraction of audio features i.e., speech processing and the classification of the music data in the servers to increase the accessibility for the users. The first stage involves in the extraction of multiple audio features [2]. The characteristics extracted from the training data are used to build a classifier in the second stage. There have been a variety of techniques to categorizing music into different genres. With so much music data available on the web, the organizations require an automatic music genre a classification system. Various types of feature extraction are used in each implementation. Some studies use pitch, timber, and beat as classification criteria. Deep learning algorithms are a popular classification technique that is used to train a large database [1].

Using Artificial Neural Networks and K-Nearest Neighbour technique, we offer a unique approach for automatic music genre classification. Music genre classification forms a basic step for building a strong recommendation system. The idea behind this project is to see how to handle sound files in python, compute sound and audio features from them, run Machine Learning Algorithms on them, and see the results. In a more systematic way, the main aim is to build and train a machine learning model, which serves the need of classifying music samples into their respective genres. It aims to predict the genre using an audio file as its input. The objective of automating the music classification is to make the selection of songs quick and less cumbersome. If one has to manually classify the songs or music, one has to listen to a whole lot of songs and then select the genre. This is not only time-consuming but

also, difficult. Automating music genre classification can be optimized to help to find and tag audio files with their genres, and artists easily [3].

1.2 Objectives of the project

- This project classifies the music genre dataset and identifies the genre of the music. The dataset consists of 6 different Indian music. We are using ann, knn, svm, decision tree for classification. The most accurate result is given as the genre of the selected music
- The deep learning approach that is used is ANN. This model is trained end to end to predict the genre label of an audio signal solely using its spectrogram.
- It utilizes hand-crafted features, both from the time domain and frequency domain.
- We train three traditional machine learning classifiers with these features and compare their performance.
- The features that contribute the most towards this classification task are identified.

1.2 Scope of the project

- Our project is directed towards developing a classification system of music which can predict the different categories of genre of music. We are using an automated system that identifies different types of genres according to the direct flow, tempo of music, and the beat and rhythm from the given audio inputs. we are doing this by using the Indian music dataset.

- The main aim is to build and train a machine learning model, which serves the need of classifying music samples into their respective genres. This model is aimed towards providing a standard system that can access accurate musical data and is equipped with tool extracting features and audio content that helps in reaching a good accuracy level by classifying music into its suitable genre appropriately. It can also identify music according to the style and features of music. It aims to predict the genre using an audio file as its input. The objective of automating the music classification is to make the selection of songs quick and less cumbersome. If one has to manually classify the songs or music, one has to listen to a whole lot of songs and then select the genre. This is not only time-consuming but also difficult. Automating music genre classification can be optimized to help find and tag audio files with their genres, and artists easily.

- Since the field of application of voice intelligence is vast and reasonably new, the forthcoming research in this domain is quite tremendous and our music genre classification system can also be developed into a standalone application which can be installed and utilized by websites to automatically classify the audio inputs that have been given to it.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

1.3 Advantages

- Ann layers are reduced to 9.
- Its supervised, hence less chances of errors
- Simple to understand
- Simple to execute
- User friendly
- Can be used for multiple platforms
- simplest machine learning and deep learning techniques used to classify the music genre.

1.2 Disadvantages

- multiple classifiers are required
- Difficult to understand and interpret the final model, variables weights and individual's impact.
- Computation cost is quite high because we need to compute distance of each query instance to all training samples. Some indexing (e.g., K-D tree) may reduce this computational cost.
- Logistic regression attempts to predict outcomes based on a set of independent variables, but if researchers include the wrong independent variables, the model will have little to no predictive value.
- Choosing a "good" kernel function is not easy
- Long training time on large data sets.

International Journal for Innovative
Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

1.6 Applications

- To categorize and organize songs is based on the genre, which is identified by some characteristics of the music such as rhythmic structure, harmonic content and instrumentation.

- Being able to automatically classify and provide tags to the music present in a user's library, based on genre, would be beneficial for audio streaming services such as Spotify and iTunes.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

1.7 Hardware and Software Requirements

H/W Configuration:

• Processor	I3/Intel Processor
• Hard Disk	160GB
• Key Board	Standard Windows Keyboard
• Mouse	Two or Three Button Mouse
• Monitor	SVGA
• RAM	8Gb

S/W Configuration:

• Operating System	Windows 10
• Server-side Script	Mysql-sqlyog
• IDE	Pycharm
• Libraries Used	Matplotlib, sklearn, pydub, mlp, pyqt5
• Technology	Python 3.6+

CHAPTER 2

LITERATURE SURVEY

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

2.1 Existing System

- This project presents an application that uses Machine Learning techniques to do Music Genre Classification.
- The Labrosa library, a Python package for music and audio analysis, was used. It outlines the basic processes for developing music information retrieval systems.
- The dataset used in our project is Indian Music Genre dataset that encompasses different Indian music genres in the form of mp3 files. The dataset was collected for just the purpose of using it to train Machine learning models into classifying Indian music genres
- MFCC was used for feature extraction. By using frequency and decibel scale, instead of amplitude (loudness), we can see a much clear spectrogram. KNN and CNN algorithms were implemented to train the model and perform classification on the dataset [14][15][16][24][41].
- The Python based Labrosa has been employed to extract the audio features and plot the spectrograms for further training [19].
- The Adam optimizer was used after evaluating other optimizers because it gave the best results [23][35].

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

2.2 Proposed System

- An accurate way of representing and identifying music label is based on musical features and characteristics. The accuracies can be affected by the quality and the vastness of the dataset [22].
- To achieve a prediction with decent accuracy the dataset needs to encompass a great variety of genres with large number of tracks for the model to train on.
- The problem we overcame during the project was lack of a proper classification model. And also, the problem of improper classification of same type of music.
- This system easily provides same genres of songs of an artist using machine learning algorithms. So that, the music listener and clients no longer need to search songs of an artist and genre for a longer time.
- Fast Fourier Transformations (FFT) is the major technique used for converting audio file [7][40].
- Artificial Neural Network algorithm was implemented for training the model and performing classification on the dataset [6].

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

CHAPTER 3

PROPOSED ARCHITECTURE

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

ARCHITECTURE

In this project, we propose a hybrid architecture which consists of the paralleling KNN and ANN. They focus on spatial features and temporal frame orders extraction and feature extraction. Then the two outputs are fed into KNN algorithm which finds K number of nearest neighbor based on features and maximum neighbors belonging to particular class and gives an output [30]. Then, the dataset is fed into the training model for classification. After the model has been trained, the imported audio file is classified [21] fed to the model and the music genre is predicted as output. We are using an automated system that identifies different types of genres according to the direct flow, tempo of music, and the beat and rhythm from the given audio inputs [7][8][17][31][32]. This is done by using the Indian music dataset. The algorithms that were implemented are KNN Algorithm to find the nearest neighbor and ANN classifier to classify the music genre of the imported audio file [4][5][6][37][38][39].

- **KNN**

  K-Nearest Neighbors is one of the most fundamental machine learning approaches. Despite its simplicity in idea, KNN is a complicated algorithm that produces reasonably good accuracy on most jobs. KNN tries out various K values to see which one produces the best results. KNN is a supervised learning algorithm, which means that the samples in the dataset must have labels/classes assigned to them. KNN is a non-parametric method, for starters. This means that when the model is utilized, no assumptions about the dataset are made. Rather, the model is built solely from the data provided. Second, when employing KNN, the dataset is not separated into training and test sets. When the model is asked to generate predictions, KNN makes no generalizations between the training and testing sets, therefore all of the training data is used [44].

- **ANN**

  Artificial Neural networks are one method for classifying data. Because Neural Networks [29][27][33][36] require some form of picture representation, the audio samples were converted to Mel Spectrograms to achieve this. To do so, the low-level characteristics in time and frequency for audio files (.wav) that were contained in the dataset were extracted and mapped these low-level features into a .json file and then this was used to train the model. Feature extraction was done using Fast Fourier Transform (FFT). Any incoming audio file was then identified on the basis of its low-level features and comparing these with the low-level features of audio files trained in the model and then the prediction of genre to which the new audio file belongs to was made. For the ANN model, we had used the Adam optimizer for training the

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

model. The epoch that was chosen for the training model is 100. The Adam optimizer was chosen because it gave the best results after evaluating other optimizers. The model accuracy can be increased by further increasing the epochs but after a certain period we may achieve a threshold, so the value should be determined accordingly.

- **SVM**

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning. The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane. SVM chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called as support vectors, and hence algorithm is termed as Support Vector Machine. SVM algorithm can be used for **Face detection, image classification, text categorization,** etc. [11][12][18][20][34].

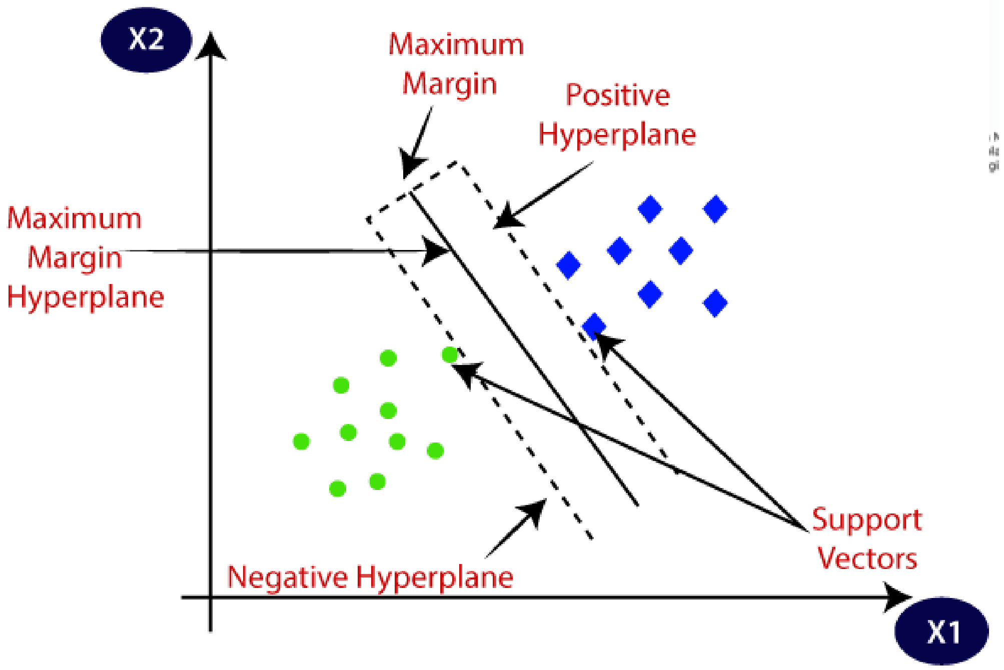
SVM can be of two types:

- **Linear SVM:** Linear SVM is used for linearly separable data, which means if a dataset can be classified into two classes by using a single straight line, then such data is termed as linearly separable data, and classifier is used called as Linear SVM classifier.
- **Non-linear SVM:** Non-Linear SVM is used for non-linearly separated data, which means if a dataset cannot be classified by using a straight line, then such data is termed as non-linear data and classifier used is called as Non-linear SVM classifier.

Consider the below diagram in which there are two different categories that are classified using a decision boundary or hyperplane:

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Fig:3.1



The advantages of support vector machines are:

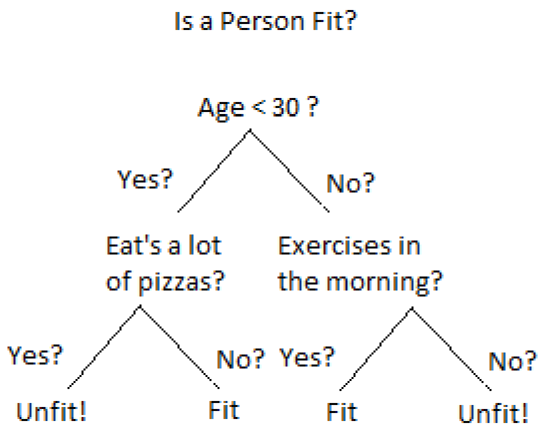
- Effective in high dimensional spaces.
- Still effective in cases where number of dimensions is greater than the number of samples.
- Uses a subset of training points in the decision function (called support vectors), so it is also memory efficient.
- Versatile: different Kernel functions can be specified for the decision function. Common kernels are provided, but it is also possible to specify custom kernels.

- **Decision Tree**

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Decision Trees are a type of Supervised Machine Learning (that is you explain what the input is and what the corresponding output is in the training data) where the data is continuously split according to a certain parameter. The tree can be explained by two entities, namely decision nodes and leaves. The leaves are the decisions or the final outcomes. And the decision nodes are where the data is split [24].

Fig:3.2



A decision tree is drawn upside down with its root at the top. In the image on the left, the bold text in black represents a condition/internal node, based on which the tree splits into branches/ edges. The end of the branch that doesn't split anymore is the decision/leaf, in this case, whether the passenger died or survived, represented as red and green text respectively.

Although, a real dataset will have a lot more features and this will just be a branch in a much bigger tree, but you can't ignore the simplicity of this algorithm. The feature importance is clear and relations can be viewed easily. This methodology is more commonly known as learning decision tree from data and above tree is called Classification tree as the target is to classify passenger as survived or died. Regression trees are represented in the same manner, just they predict continuous values like price of a house. In general, Decision Tree algorithms are referred to as CART or Classification and Regression Trees [45].

- **Logistic Regression**

  Logistic regression is a supervised learning classification algorithm used to predict the probability of a target variable. The nature of target or dependent variable is dichotomous, which means there would be only two possible classes. In simple words, the dependent variable is binary in nature having data coded as either 1 (stands for success/yes) or 0 (stands for failure/no). Mathematically, a logistic regression model predicts P(Y=1) as a function of X. It is one of the simplest ML algorithms that can be used for various classification problems such as spam detection, Diabetes prediction, cancer detection etc.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

CHAPTER 4
IMPLEMENTATION

4.1 Algorithm

1. Import the audio file[42]

2. Define a function to get the distance between feature vectors and find neighbors [9][10][25]

3. Identify the nearest neighbors using KNN algorithm

4. Define a function for model evaluation

5. Extract features from the dataset [43].

6. Train and test split on the dataset

7. Make prediction using KNN algorithm and ANN classifier to get the accuracy on test data

8. After the model has been trained, predict the genre of the audio file by selecting an audio file from the library

9. The genre of the audio file is predicted by the training model

10. The classified music genre is displayed as output.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

4.2 Code

Fig 4.2.1 Code of main.py

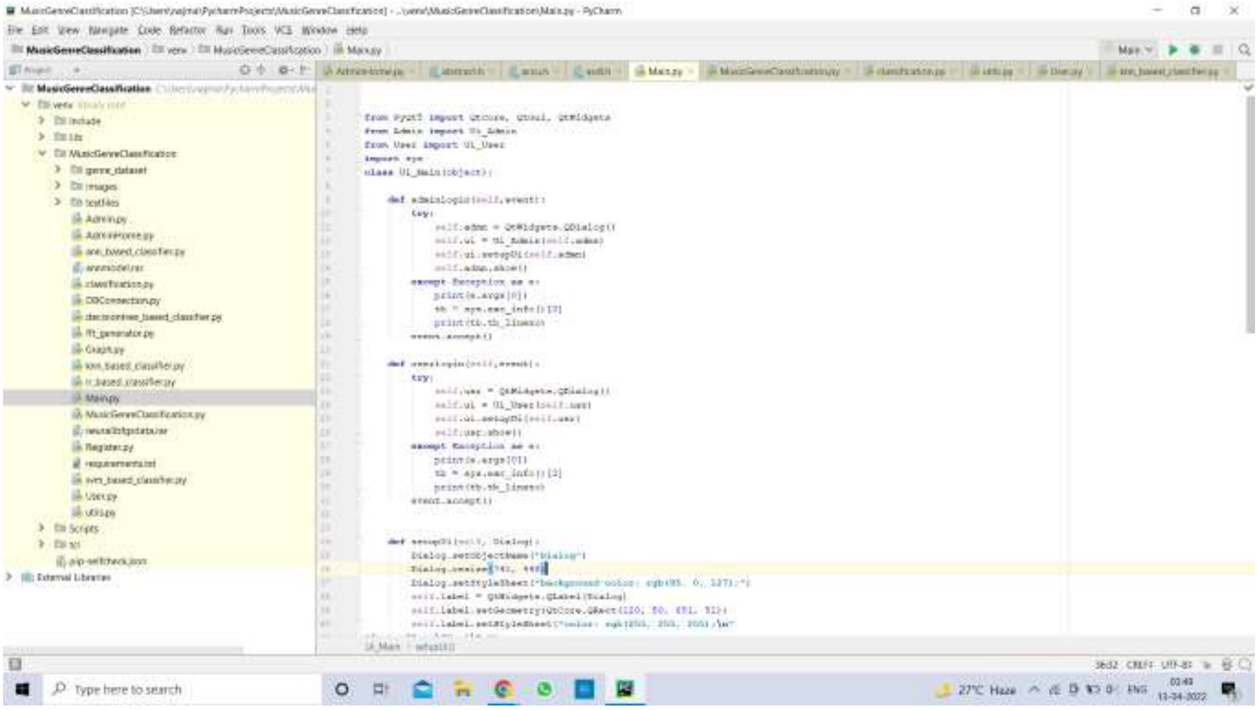
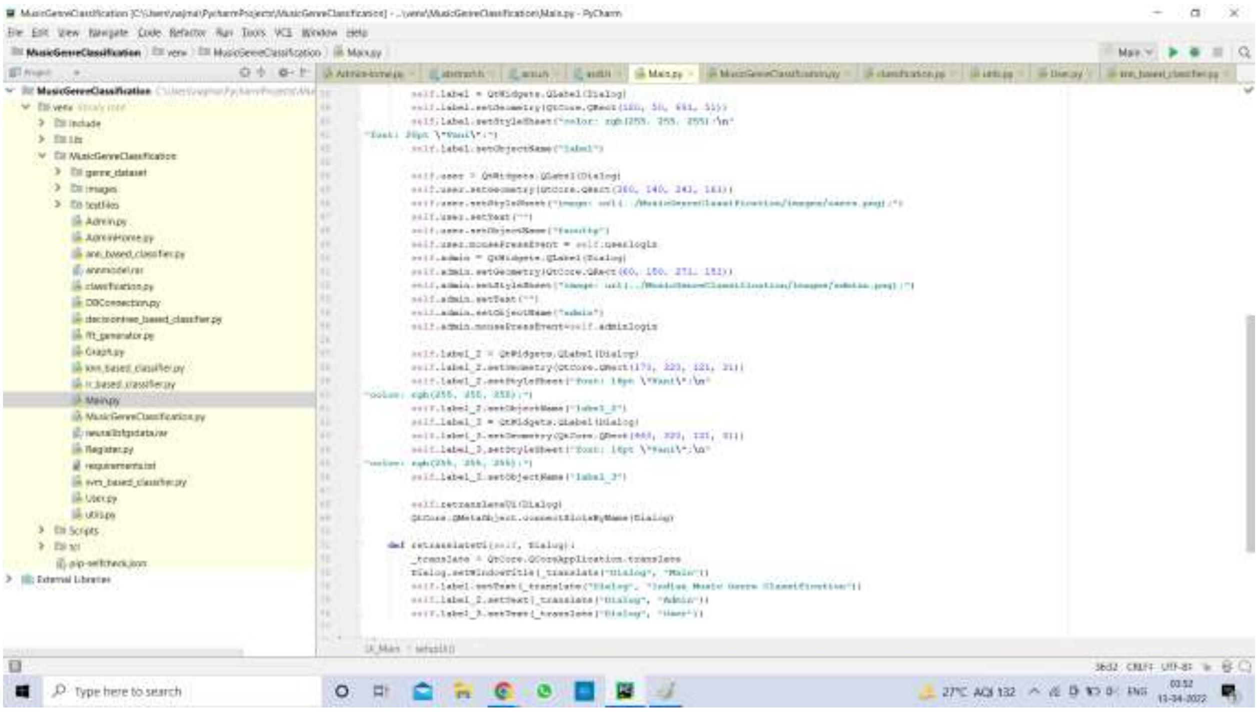
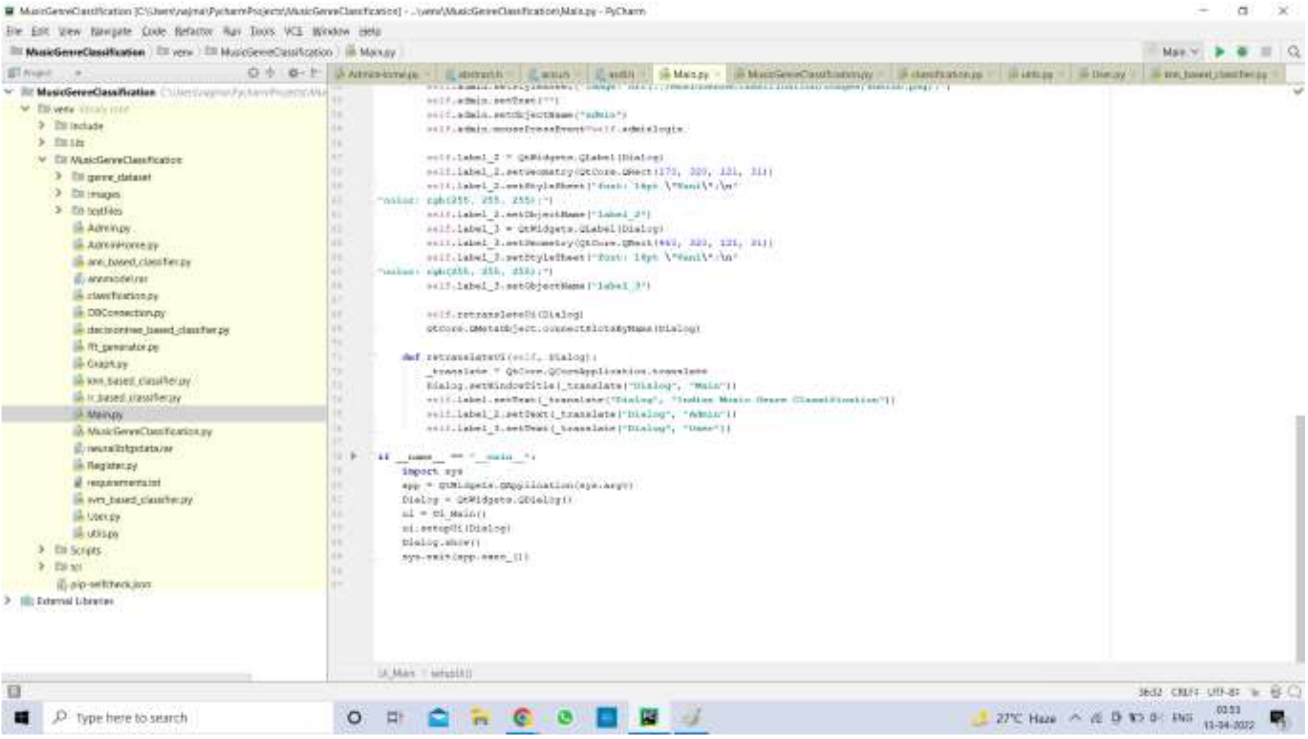


Fig 4.2.2



International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Fig 4.2.3



CHAPTER 5

RESULTS

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Fig 5.1 Main page

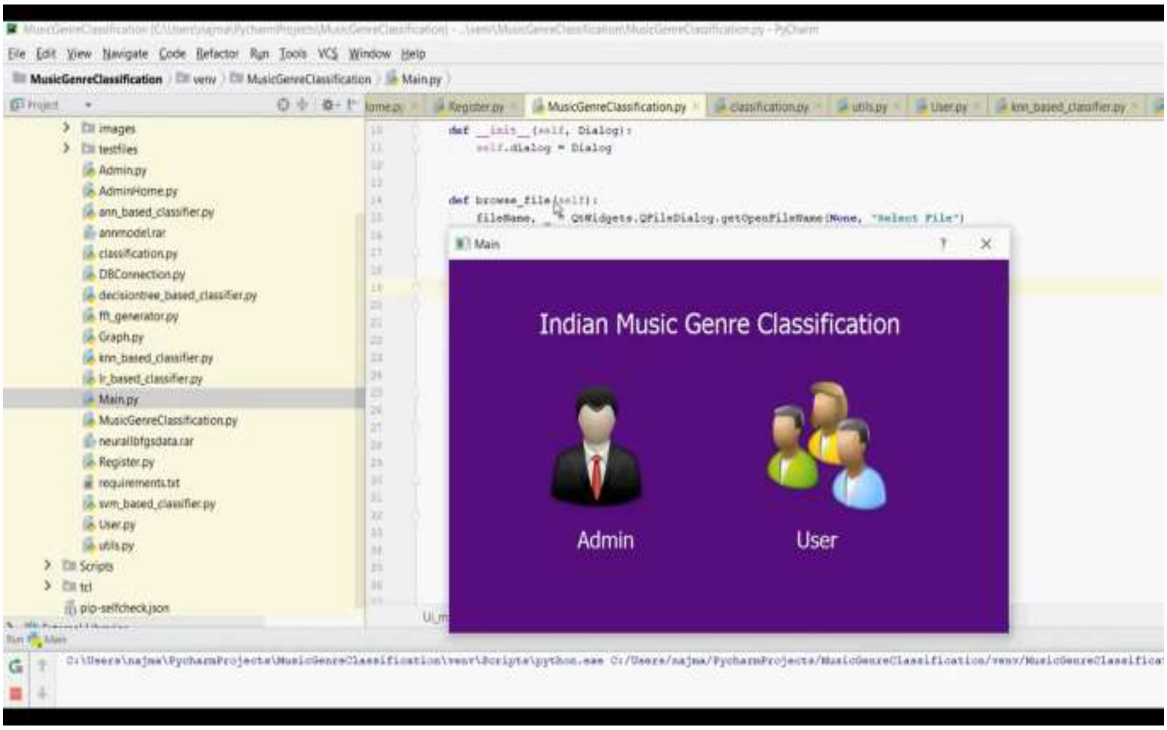
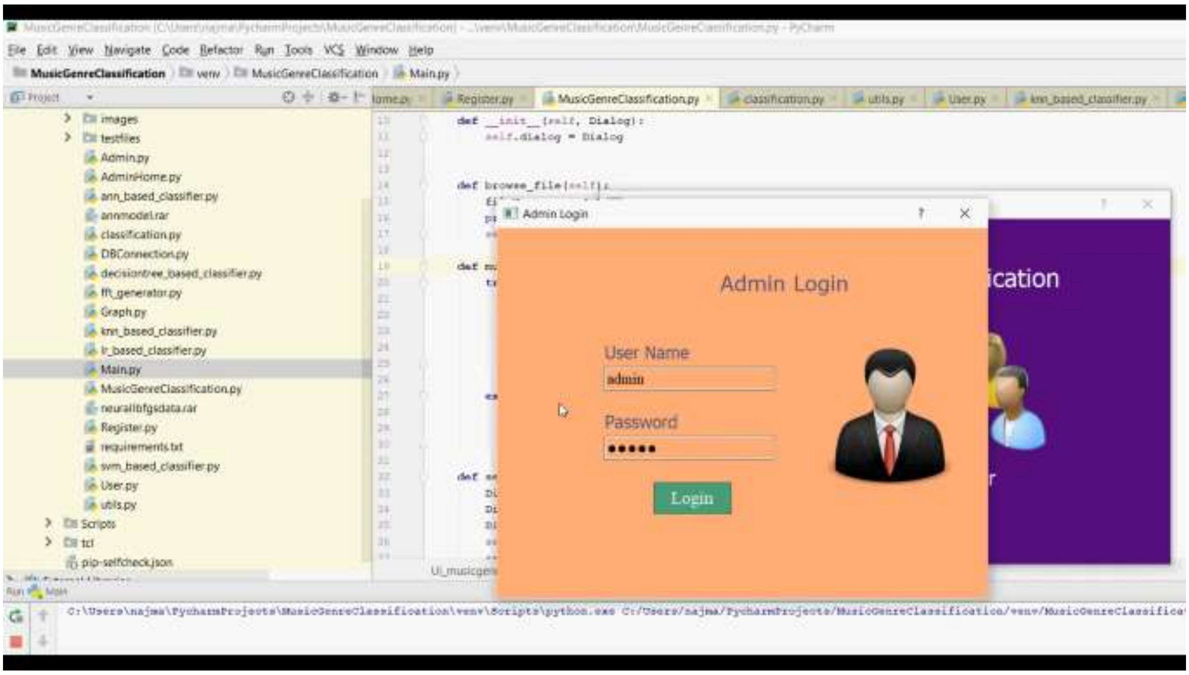


Fig 5.2 Admin Login



International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Fig 5.3 Admin Processing the Dataset

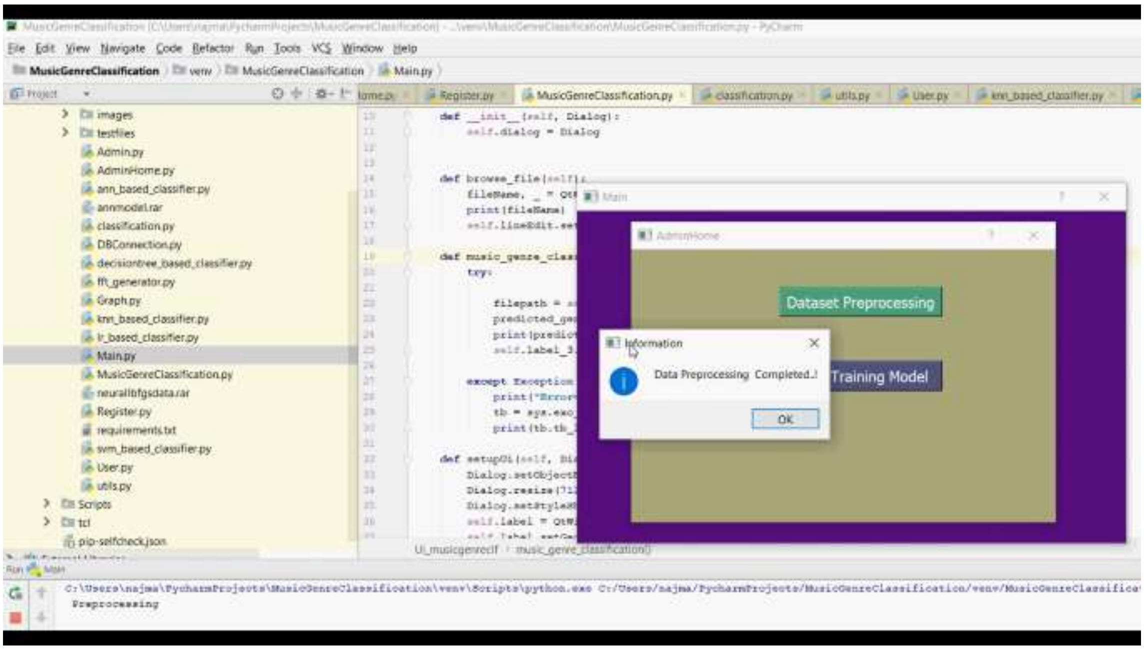


Fig 5.4 Training the model

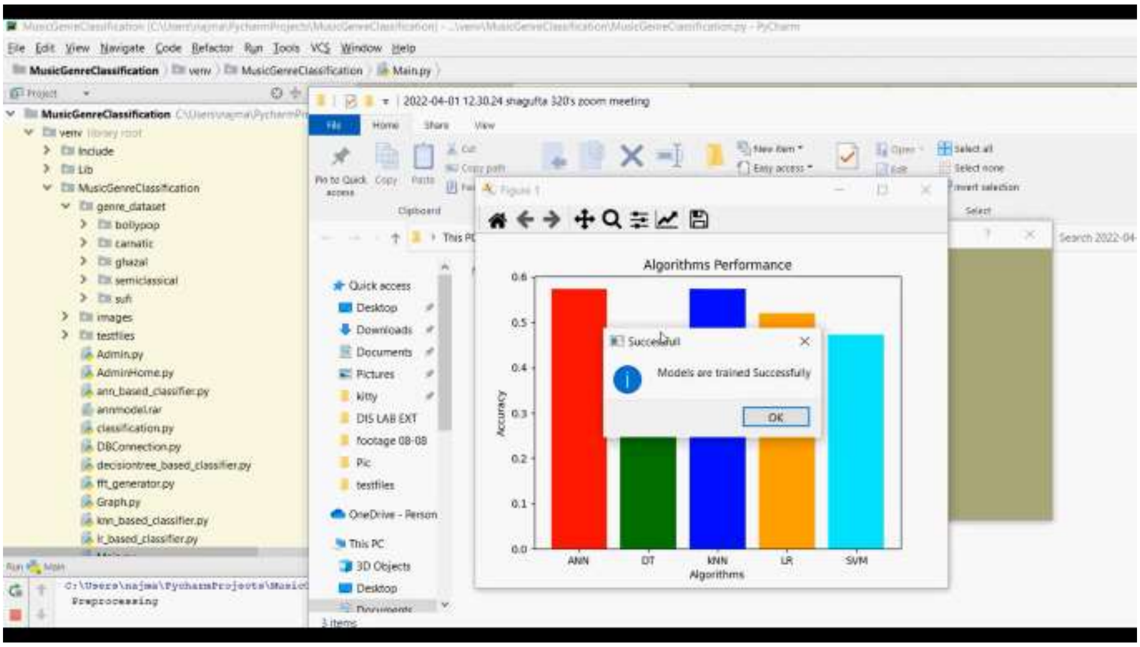


Fig 5.5 User registration

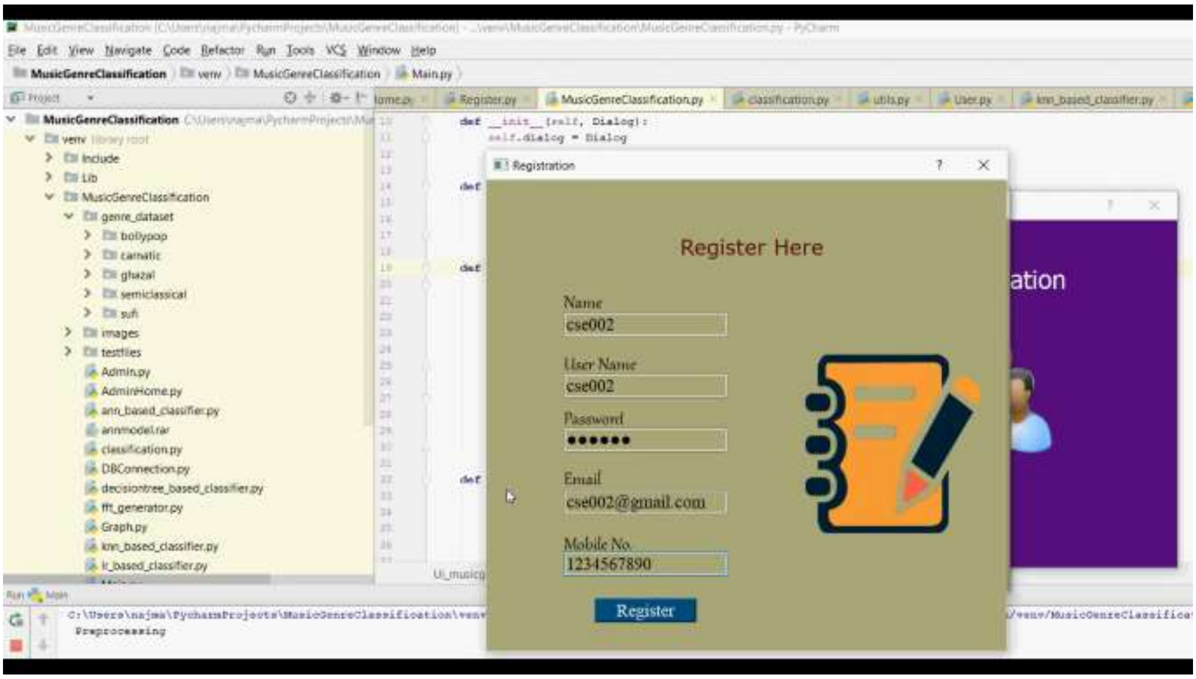


Fig 5.6 User Login

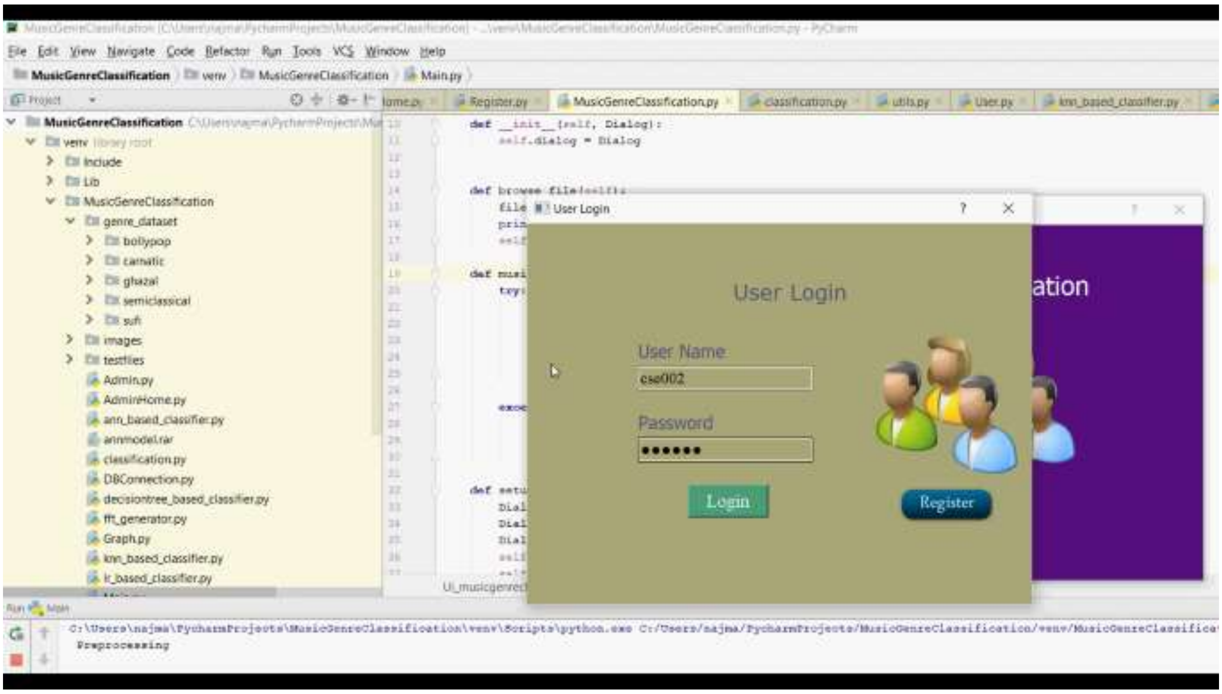


Fig 5.7 Selecting audio file

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

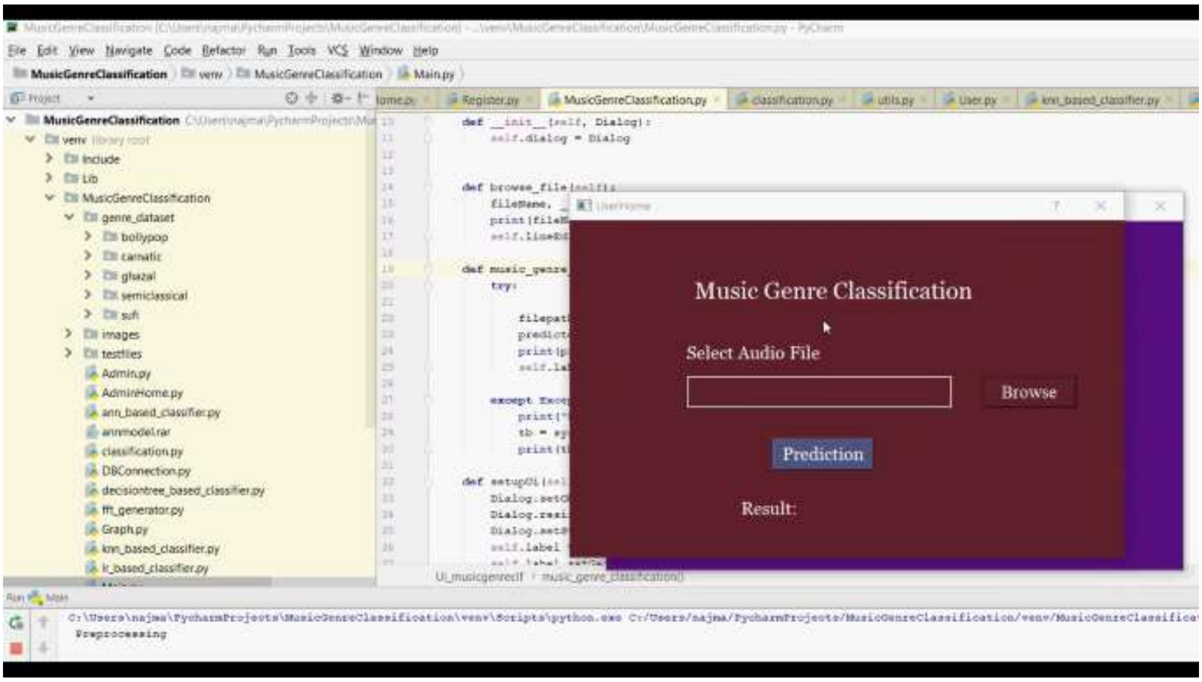
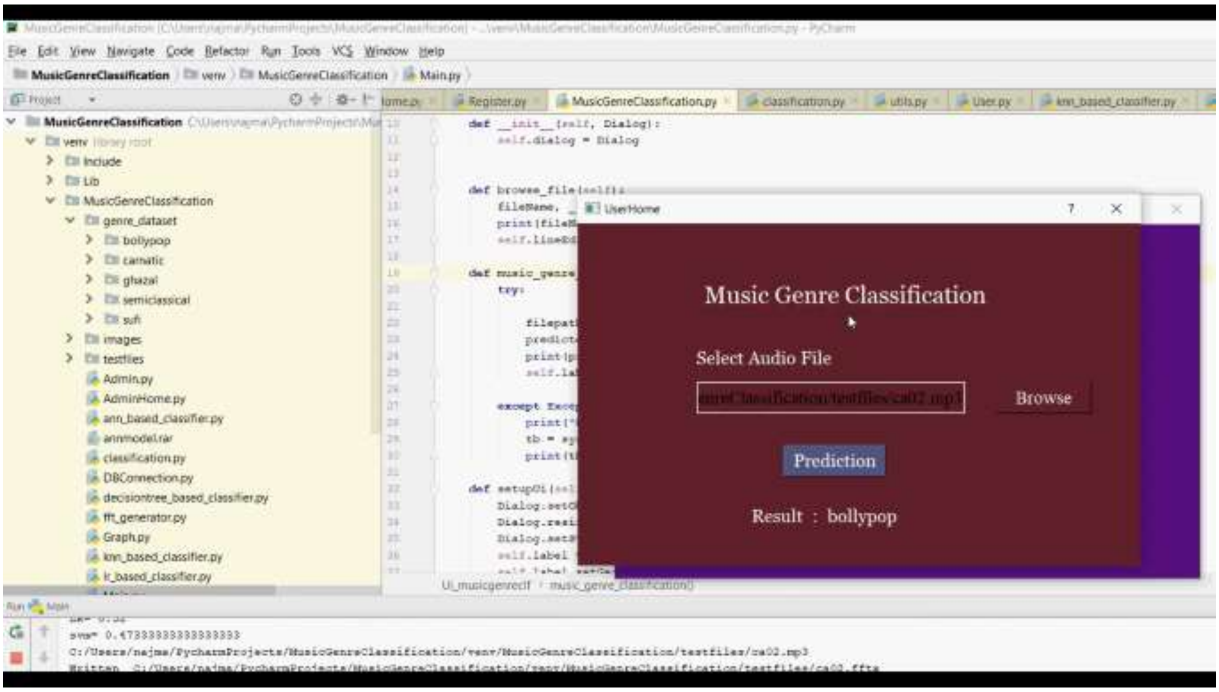


Fig 5.8 Result (1)



International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Fig 5.9 Result (2)

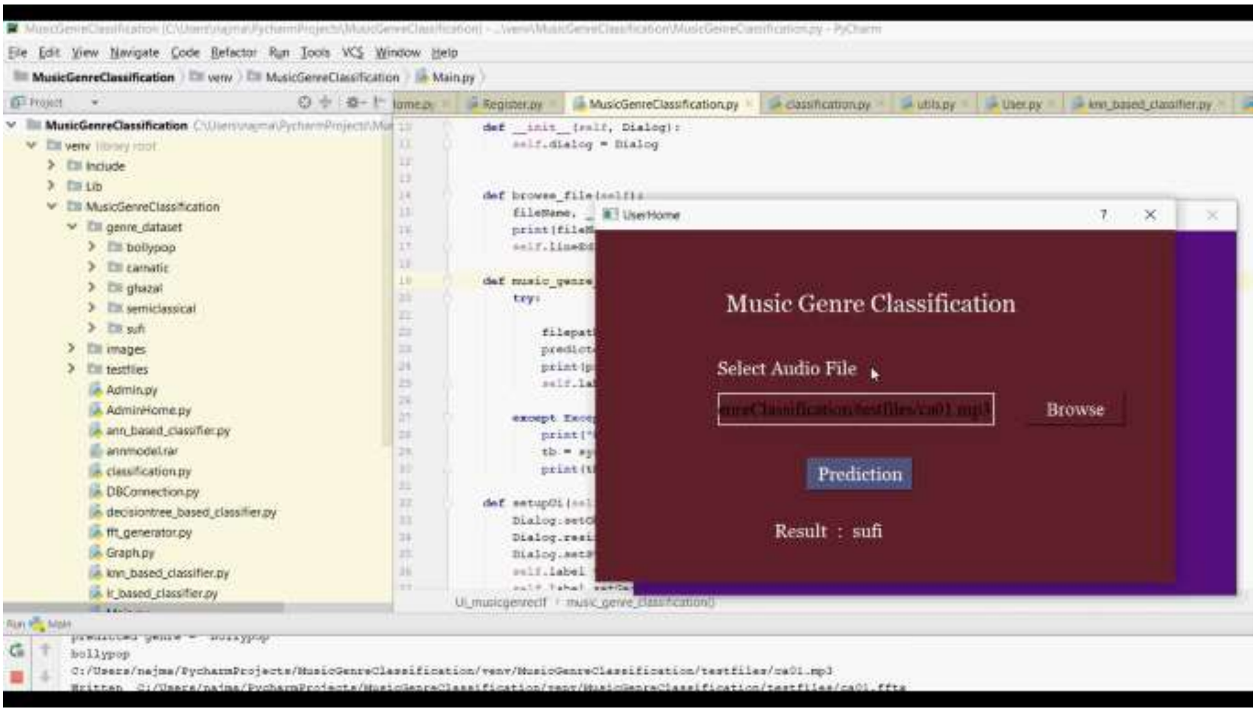
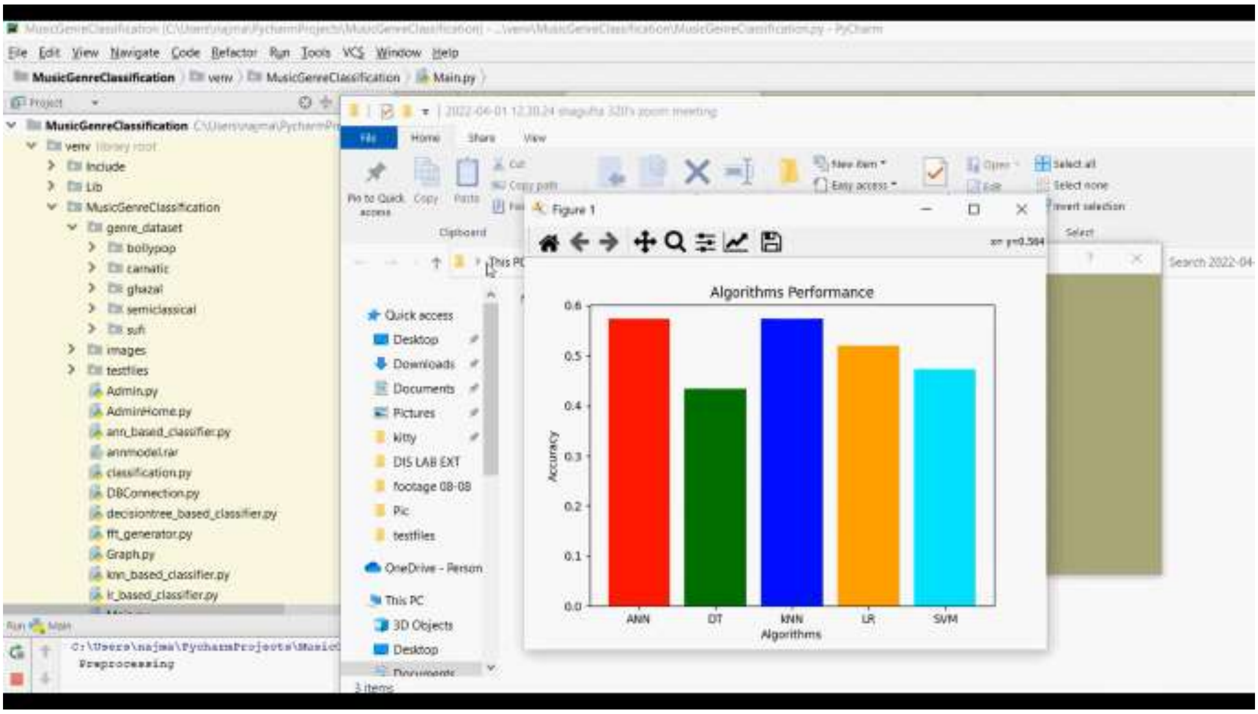


Fig 5.10 Accuracy Graph



International Journal for Innovative
Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

CHAPTER 6

CONCLUSION

CONCLUSION

- This project studies several methods of music genre classification. We study several audio feature extraction methods using digital signal processing methods, such as MFCC, FFT etc. [26] And we propose two main architectures for music genre classification. One is the combination with different level features with decision tree model, and the other one is the combination with a series of FFT with KNN algorithm and ANN classifier. We quantitatively compare the performance of classifying Sufi, Classical, Bollypop, Carnatic using two architectures, and find that the ANN model gives a better result with 90% accuracy [28].
- Generally speaking, in this project, we propose two architectures for music genre classification, based on digital signal processing techniques and machine learning methods. Both architectures give descent results with more than 80% whole accuracy. We have started the project with the initial setup and used FFT to extract features from audio files. After that, we use KNN classifier that finds K number of nearest neighbor based on features and maximum neighbors belonging to particular class gives as an output. We got approximately 70 per cent accuracy on the model.
- The main thing to identify and divide the audio into different features is amplitude and frequency that changes within a short span of time.
- We can visualize the audio frequency wave of amplitude and frequency with respect to time in form of a wave plot that can be easily plotted using librosa [19].

CHAPTER 7

FUTURE SCOPE

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

SCOPE

An automated system to classify the song based on the direct flow and tempo of music and to compare and draw the necessary conclusions regarding the accuracy of this Machine Learning model and the pre-existing models.

This project aims at developing a classification of music by a detailed study approach which can correctly predict

the category of the Indian Classical Music Genre and its confidence level.

To produce quality system product accessing accurate musical data and availability of tool extracting features and audio content and reach good accuracy level by classifying model of music into its connective genre correctly and identify music according to its style and features, rather than manually entering the genre and provide an overview of the developed framework for music genre classification.

This project could also be developed into a standalone application deployable on websites to automatically classify the fed audio or video inputs. The future work in this domain is quite immense since the field of application of voice intelligence is wide and fairly new.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

CHAPTER 8

REFERENCES

References

[1] A Chun. 2010. Automatic musical pattern feature extraction using convolutional neural network. In Proc. Int. Conf. Data Mining and Applications.

[2] Aaron Van Den Oord 2016. Wavenet: A generative model for raw audio. arXiv preprint arXiv:1609.03499 .

[3] Abdel-rahman Mohamed 2014. Convolutional neural networks for speech recognition. IEEE/ACM Transactions on audio, speech, and language processing 22(10):1533–1545.

[4] Alex Krizhevsky 2012. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems. pages 1097–1105.

[5] Alex Krizhevsky 2014. Dropout: A simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research 15(1):1929–1958.

[6] Alexander Schindler. 2016. Parallel convolutional neural networks for music genre and mood classification. MIREX2016 .

[7] Alex Waibel. 1998. Recognition of music types. In Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on. IEEE, volume 2, pages 1137–1140.

[8] Alessandra Lumini 2016. Combining visual and acoustic features for music genre classification. Expert Systems with Applications 45:108–117.

[9] Alex Graves 2016. Wavenet: A generative model for raw audio. arXiv preprint arXiv:1609.03499 .

[10] Andrew Senior 2016 A generative model for raw audio. arXiv preprint arXiv:1609.03499 .

[11] Andrew Y Ng. 2004. Feature selection, l 1 vs. l 2 regularization, and rotational invariance. In Proceedings of the twenty-first international conference on Machine learning. ACM, page 78.

[12] Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 .

[13] Andreas Rauber. 2005. Evaluation of feature extractors and psycho-acoustic transformations for music genre classification. In ISMIR. pages 34–41.

[14] Antoni B Chan 2010. Automatic musical pattern feature extraction using convolutional neural network. In Proc. Int. Conf. Data Mining and Applications.

[15] Aren Jansen 2017. Audio set: An ontology and human-labeled dataset for audio events. In Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on. IEEE, pages 776–780.

[16] Chadawan Ittichaichareon 2012. Speech recognition using mfcc. In International Conference on Computer Graphics, Simulation and Modeling (ICGSM’2012) July. pages 28–29.

[17] Chanwoo Kim 2012. Powernormalized cepstral coefficients (pncc) for robust speech recognition. In Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on. IEEE, pages 4101–4104

[18] Corinna Cortes 1995. Supportvector networks. Machine learning 20(3):273–297.

[19] Dan Ellis. 2007. Chroma feature analysis and synthesis. Resources of Laboratory for the Recognition and Organization of Speech and Audio-LabROSA .

[20] Dan Ellis. 2005. Song-level features and support vector machines for music classification. In ISMIR. volume 2005, pages 594–599

[21] Dan-Ning Jiang 2002. Music type classification by spectral contrast feature. In Multimedia and Expo, 2002. ICME'02. Proceedings. 2002 IEEE International Conference on. IEEE, volume 1, pages 113–116. done

[22] Daniel PW Ellis 2017. Audio set: An ontology and human-labeled dataset for audio events. In Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on. IEEE, pages 776–780.

[23] Diederik P Kingma 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 .

[24] Donald Geman. 1997. Shape quantization and recognition with randomized trees. Neural computation 9(7):1545–1588.

[25] Frank Kurth. 2010. ¨ Cyclic tempograma mid-level tempo representation for musicsignals. In Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on. IEEE, pages 5522–5525. done

[26] Giorgio Zoia. 2005. On the modeling of time information for automatic genre recognition systems in audio signals. In ISMIR. pages 666–671

[27] Geoffrey Hinton 2014. Dropout: A simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research 15(1):1929–1958.

[28] George Tzanetakis 2002. Musical genre classification of audio signals. IEEE Transactions on speech and audio processing 10(5):293– 302.

[29] Gerald Penn, and. 2014. Convolutional neural networks for speech recognition. IEEE/ACM Transactions on audio, speech, and language processing 22(10):1533–1545. DONE3`11

[30] Hagen Soltau 1998. Recognition of music types. In Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on. IEEE, volume 2, pages 1137–1140.

[31] Heiga Zen 2016. Wavenet: A generative model for raw audio. arXiv preprint arXiv:1609.03499 .

[32] Hong-Jiang Zhang 2002. Music type classification by spectral contrast feature. In Multimedia and Expo, 2002. ICME'02. Proceedings. 2002 IEEE International Conference on. IEEE, volume 1, pages 113–116.

[33] Hui Jiang 2014. Convolutional neural networks for speech recognition. IEEE/ACM Transactions on audio, speech, and language processing 22(10):1533–1545.

[34] Jerome Friedman. 2001. The elements of statistical learnine.

[35] Jimmy Ba. 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 .

[36] Lonce Wyse. 2017. Audio spectrogram representations for processing with convolutional neural networks. arXiv preprint arXiv:1706.09559 .

[37] Loris Nanni 2016. Combining visual and acoustic features for music genre classification. Expert Systems with Applications 45:108–117.

[38] Martin Westphal 1998. Recognition of music types. In Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on. IEEE, volume 2, pages 1137–1140.

[39] Marvin Ritter. 2017. Audio set: An ontology and human-labeled dataset for audio events. In Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on. IEEE, pages 776–780.

[40] Richard M Stern. 2012. Powernormalized cepstral coefficients (pncc) for robust speech recognition. In Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on. IEEE, pages 4101–4104

[41] Siwat Suksri 2012. Speech recognition using mfcc. In International Conference on Computer Graphics, Simulation and Modeling (ICGSM'2012) July. pages 28–29.

[42] Steve Tjoa. 2017. Music information retrieval. https://musicinformationretrieval.com/spectral_features.html. Accessed: 2018-02-20.

[43] Tom LH Li, 2010. Automatic musical pattern feature extraction using convolutional neural network. In Proc. Int. Conf. Data Mining and Applications.

[44] Tutorialpoint

[45] Vladimir Vapnik. 1995. Supportvector networks. Machine learning 20(3):273–297.

[46] Yali Amit and. 1997. Shape quantization and recognition with randomized trees. Neural computation 9(7):1545–1588.