St. Thomas' College of Engineering and Technology

Audio Processing using Pattern Recognition for Music Genre Classification

Prepared By

Avik Bose	181220110074 (University Roll No. 12200218052)
Sivangi Chatterjee	181220110110 (University Roll No. 12200218016)
Srishti Ganguly	181220110116 (University Roll No. 12200218010)
Hrithik Raj Prasad	181220110086 (University Roll No. 12200218040)

Under the guidance of

Dr. Arijit Ghosal

(Associate Professor, IT Dept)

Dr. Ranjit Ghoshal

(Associate Professor, IT Dept)

Project Report

Submitted in partial fulfilment of the requirement for the degree of B.Tech in Information Technology

Department of Information Technology

Affiliated to

Maulana Abul Kalam Azad University of Technology, West Bengal

June, 2022

This is to certify that the work in preparing the project entitled Music Genre Classification has been carried out by Avik Bose, Sivangi Chatterjee, Srishti Ganguly and Hrithik Raj Prasad under our guidance during the session 2021-2022 and accepted in partial fulfilment of the requirement for the degree of Bachelor of Technology in Information Technology.

____

Dr. Arindam Chakravorty

Head, Department of Information Technology

____

Dr. Arijit Ghosal

Department of Information Technology

____

Dr. Ranjit Ghoshal

Department of Information Technology

Acknowledgement

I am overwhelmed in all humbleness and gratefulness to acknowledge my depth of knowledge to everyone who has helped me to put these ideas, well above the level of simplicity and into something concrete. I would like to express my special thanks to Dr. Arijit Ghosal for his immense help in guiding us throughout. I would also like to take the opportunity to thank Dr. Ranjit Ghoshal for mentoring and guiding us. I also would like to thank my parents and friends who have motivated me and inspired me during these difficult times. The project would not have been achievable if not for the aforementioned assistance.

Avik Bose

Sivangi Chatterjee

Srishti Ganguly

Hrithik Raj Prasad

Vision and Mission of the College

Vision:

To evolve itself into an industry oriented research based recognized hub of creative solutions in various fields of engineering by establishing progressive teaching-learning process with an ultimate objective of meeting technological challenges faced by the nation and the society

Mission:

- To create opportunities for students and faculty members in acquiring professional knowledge and developing social attitudes with ethical and moral values
- To enhance the quality of engineering education through accessible comprehensive, industry and research oriented teaching-learning process
- To satisfy the ever-changing needs of the nation for evolution and absorption of sustainable and environment friendly technologies

Vision and Mission of the Department

Vision:

To promote the advancement of learning in Information Technology through research-oriented dissemination of knowledge which will lead to innovative applications of information in industry and society.

Mission:

- To incubate students, grow into industry ready professionals, proficient research scholars and enterprising entrepreneurs.
- To create a learner-centric environment that motivates the students in adopting emerging technologies of the rapidly changing information society.
- To promote social, environmental and technological responsiveness among the members of the faculty and students.

Program Educational Objectives (PEO)

Graduates of Information Technology Program shall:

PEO:

PEO1: Exhibit the skills and knowledge required to design, develop and implement IT solutions for real life problems.

PEO2: Excel in professional career, higher education and research.

PEO3: Demonstrate professionalism, entrepreneurship, ethical behaviour, communication skills and collaborative team work to adapt the emerging trends by engaging in lifelong learning.

Program Outcomes(POs)

PO: Project Mapping with Program Outcomes

PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
3	3	2	3	3	3	2	3	2	3	3	3

Correlation levels 1, 2 or 3 as defined below:

1: Slight (Low) **2: Moderate (Medium)** **3: Substantial (High)**

Justification:

Engineering Graduates will be able to:

1. **Engineering knowledge**: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialisation to the solution of complex engineering problems.

2. **Problem analysis**: Identify, formulate, review research literature, and analyse complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

3. **Design/development of solutions**: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

4. **Conduct investigations of complex problems**: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

5. **Modern tool usage**: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modelling to complex engineering activities with an understanding of the limitations.

6. **The engineer and society**: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

7. **Environment and sustainability**: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

8. **Ethics**: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

9. **Individual and team work**: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

10. **Communication**: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

11. **Project management and finance**: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

12. **Life-long learning**: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

PSO: Project Mapping with Program Specific outcomes

PSO1	PSO2	PSO3
3	-	3

Correlation levels 1, 2 or 3 are as defined below:

1: Slight (Low) **2: Moderate (Medium)** **3: Substantial (High)**

Justification:

1. This project requires me to apply programming knowledge to build an efficient and effective solution of the problem with an error free, well documented and reusable code, user friendly interface and well-organised database. Hence, the project substantially satisfies PSO1.
2. Creation of multimedia enabled web solutions using information in different forms for business, education and the society at large isn't applicable in this project, hence, PSO2 stands irrelevant/ not applicable.
3. Understanding and analysing a big complex problem and decomposing it into relatively smaller and independent modules algorithmically is done in this project, hence, PSO3 is satisfied substantially.

Index Page

Contents **Page No.**

1. Introduction

1.1 Problem Statement:

Music Genre Classification using Machine Learning.

1.2 Problem Definition:

The different genres of music are identified based on their characteristic features with the help of different classification techniques.

1.3 Objective:

The aim of this project is to build a model which correctly recognizes the genre of the music based on their characteristic audio features using Machine Learning.

1.4 Literature Survey:

Matthew Creme, Charles Burlin, Raphael Lenain, "Music Genre Classification ", Stanford University, December 15, 2016. [1]
Creme, Burlin, Lenain have presented successive steps towards building different classifying methods to identify a specific genre from an initial audio file. The final output of the algorithms used is the prediction of the genre of each input. It also quickly diverges into composer classification for classical music. Finally, the results are presented that they have obtained while studying this problem.

Vishnupriya S, and K.Meenakshi, "Automatic Music Genre Classification using Convolution Neural Network", IEEE Conference 2018. [2]
Vishnupriya and Meenakshi have researched in their paper that most of the current music genre classification techniques use machine learning techniques. A music dataset is presented which includes ten different genres. A Deep Learning approach is used in order to train and classify the system. Here, convolution neural network is used for training and classification.The accuracy level of our system is around 76%.

Arsh Chowdhry, "Music Genre Classification using CNN", May 7, 2021. [3]
Chowdhury has presented the idea behind this project which 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 create a machine learning model, which classifies music samples into different genres. It aims to predict the genre using an audio signal as its input.

Gautam Chettiar, Kalaivani S, " Music Genre Classification Techniques" Conference paper, Vellore Institute of Technology, 11 November, 2021. [4]
Chettiar and Kalaivani aim to chart out various methods and parameters essential in the classification process, with use of Deep Learning techniques and an application of a Non-Linear Frequency Cepstrum. Music genres will be classified by taking FFT coefficients, followed by thy MFCC's and both coefficients will be used as inputs for Deep Learning models such CNN, RNN, KNN, Naïve Bayes Classifier and SVM, followed by a tabulation of the obtained results.

Parul Pandey, "Music Genre Classification with Python", December 13, 2018. [5]
Pandey has compared the performance of the two classes of models in this study. Deep learning approach is the first approach and for the purpose of predicting the genre label of a signal, a CNN model would be trained usually from end-to-end, a spectrogram is used to carry out this process. Hand-crafted features were used in the second approach which was also the last approach. Four out of traditional machine learning classifiers are trained beside these features and their performance is compared afterwards.

Introduction to Machine Learning by Alex Smola and S.V.N. Vishwanathan. [6]
Smola and Vishwanathan's purpose in this paper is to provide the reader with an overview over the vast range of applications which have at their heart a machine learning problem and to bring some degree of order to the zoo of problems.

1.5 Tools and Platform:

This project has been done on Jupyter Notebook with the help of Python 3.
In this project Librosa library has been used to read the digital audio .wav files. Librosa is a python package for music and audio analysis. It provides the building blocks necessary to create music information retrieval systems. The package has the necessary functions to extract the features that are used for the genre classification.

Also Audacity has been used to ensure if all the audio files are in mono format.

1.6 Brief Discussion on Problem:

With the start of the 21st century, the pool of music tends to be ever-expanding. People nowadays tend to bid for a total genre of music rather than a single artist or band. Music is an ever growing pool. People tend to listen to the older ones as well as the newer trends in the music industry. Thus sorting these into different genres manually would be an impossible task at the present time. Companies nowadays use music classification either to be able to place perfect recommendations to their customers like Spotify or SoundCloud or simply as a product like shazam.Therefore determining which music falls under which category or genre is the first step of advancement to this. Therefore, Machine learning techniques and classification algorithms will be used in order to understand the pattern or trends of similar genres of music, and thus sort them categorically.

2. Concepts and Problem Analysis

This project is an application of Machine Learning. Machine Learning is the study of algorithms that gives computers the capability to learn and adapt on its own from previous experiences without being explicitly programmed and draw inferences from patterns in data. A general machine learning block diagram is illustrated in Fig 1. A typical machine learning problem starts with data acquisition from various sources followed by data pre-processing where raw data is filtered to improve its quality. The processed data is then used for extracting features and the dataset is split in two groups - one for training and the other for model evaluation. After the selection of a suitable machine learning model depending on the data, the model is trained using the training data set. This is followed by the training model being tested against the evaluation data set for model evaluation which involves calculating the accuracy and precision of the trained model. This evaluation can be further used to fine tune the model until a desired accuracy is achieved that can be deployed.

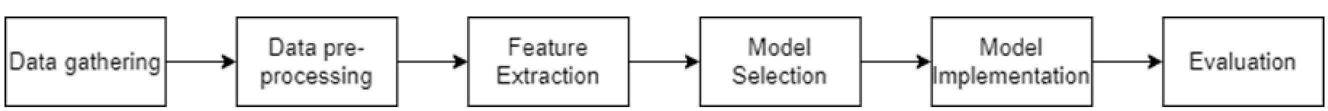

Fig. 1 : Machine Learning Block Diagram

3. Design Methodology:

3.1 Dataset:

The dataset used in this project is the GTZAN Genre Collection Dataset. It has been designed and written by George Tzanetaki. The dataset has been taken from the popular software framework MARSYAS. Marsyas (Music Analysis, Retrieval and Synthesis for Audio Signals) is an open source software framework for audio processing with specific emphasis on Music Information Retrieval applications.

Dataset consists of 1000 audio tracks each 30 seconds long. It contains 10 genres (Blues, Classical, Country, Disco, Hip-Hop, Jazz, Metal, Pop, Reggae and Rock), each represented by 100 tracks. The tracks are all 22050 Hz Mono 16-bit audio files in .wav format.

3.2 Feature Extraction:

Sound is represented in the form of an audio signal having parameters such as frequency, bandwidth, decibel etc. A typical audio signal can be expressed as a function of Amplitude and Time. The sound excerpts are digital audio files in .wav format. A typical audio processing process involves the extraction of acoustics features relevant to the task at hand, followed by decision-making schemes that involve detection, classification, and knowledge fusion. Thankfully we have some useful python libraries like librosa which make this task easier.

3.3 Feature Selection:

Each and every audio signal consists of a load of features that can be extracted. However the characteristics that are relevant to the problem have been only considered for the extraction. This removes overfitting of the data, improves accuracy and also reduces training time. The major features used are described below.

Zero Crossing Rate (ZCR):

ZCR is the rate of sign-changes along a signal, i.e., the rate at which the signal changes from positive to negative, or vice-versa. The ZCR values are in the range of 0-1. The values are related to frequencies. High ZCR value corresponds to a high frequency signal portion and vice-versa. ZCR is a good feature to discriminate among audio-silence/noise, speech/noise. It usually has higher values for metal and rock genres. So, ZCR has been considered in this project. Fig.2 shows the variation of ZCR for the mentioned genres. It is clearly visible that the graph plots differ. The X axis denotes the time and the Y axis denotes the amplitude.

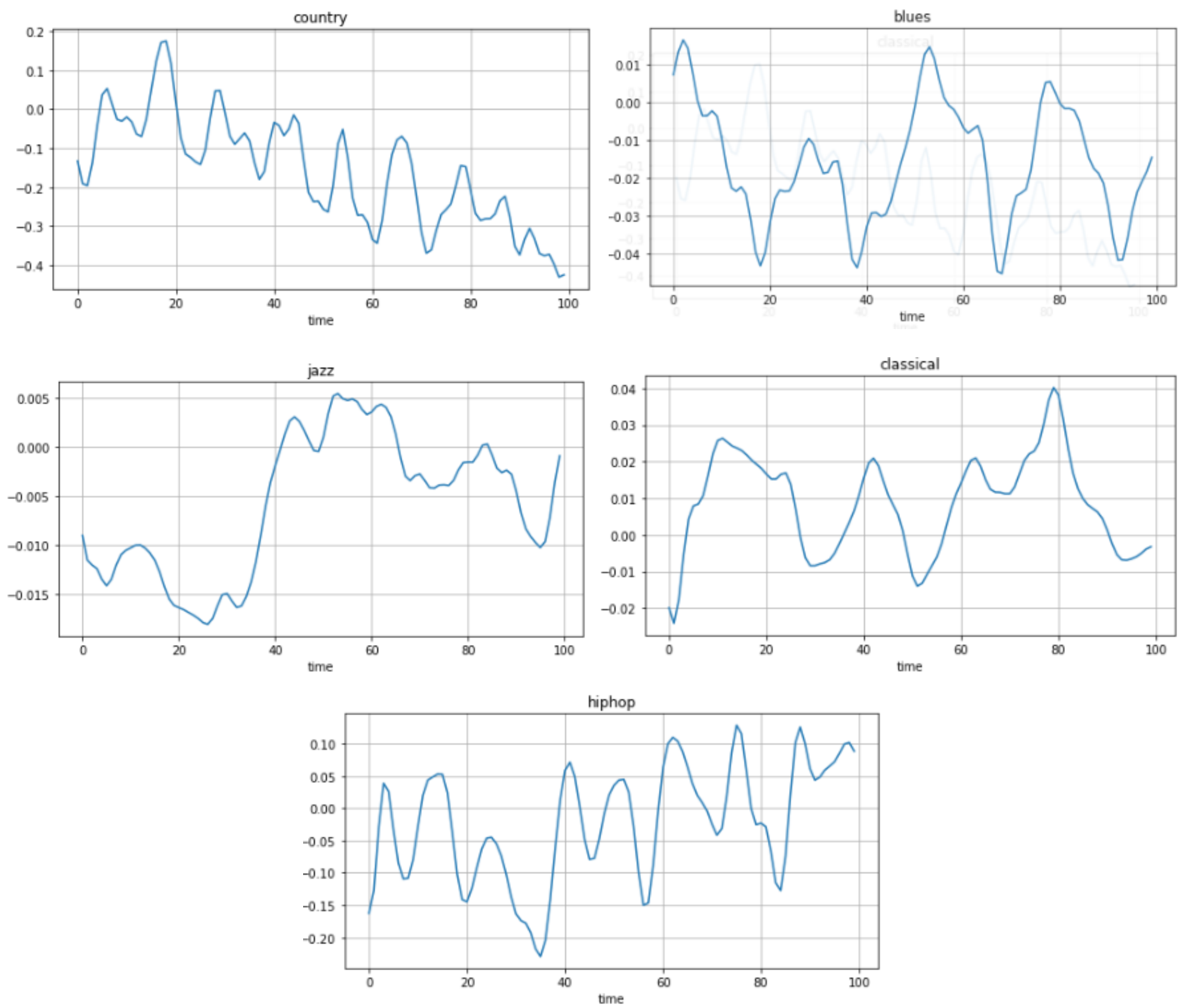

Fig 2: ZCR variation of the five genres

Spectral Centroid:

It indicates where the "centre of mass" for a sound is located and is calculated as the weighted mean of the frequencies present in the sound. If two songs, one from a blues genre and the other belonging to metal are compared, the blues genre song which is the same throughout its length, will have the spectral centroid lie somewhere near the middle of its spectrum while that for a metal song would be towards either of its ends. Fig.3 shows the Spectral Centroid and its plot variation for the mentioned genres. The

X axis denotes the time and the Y axis denotes the centroid.

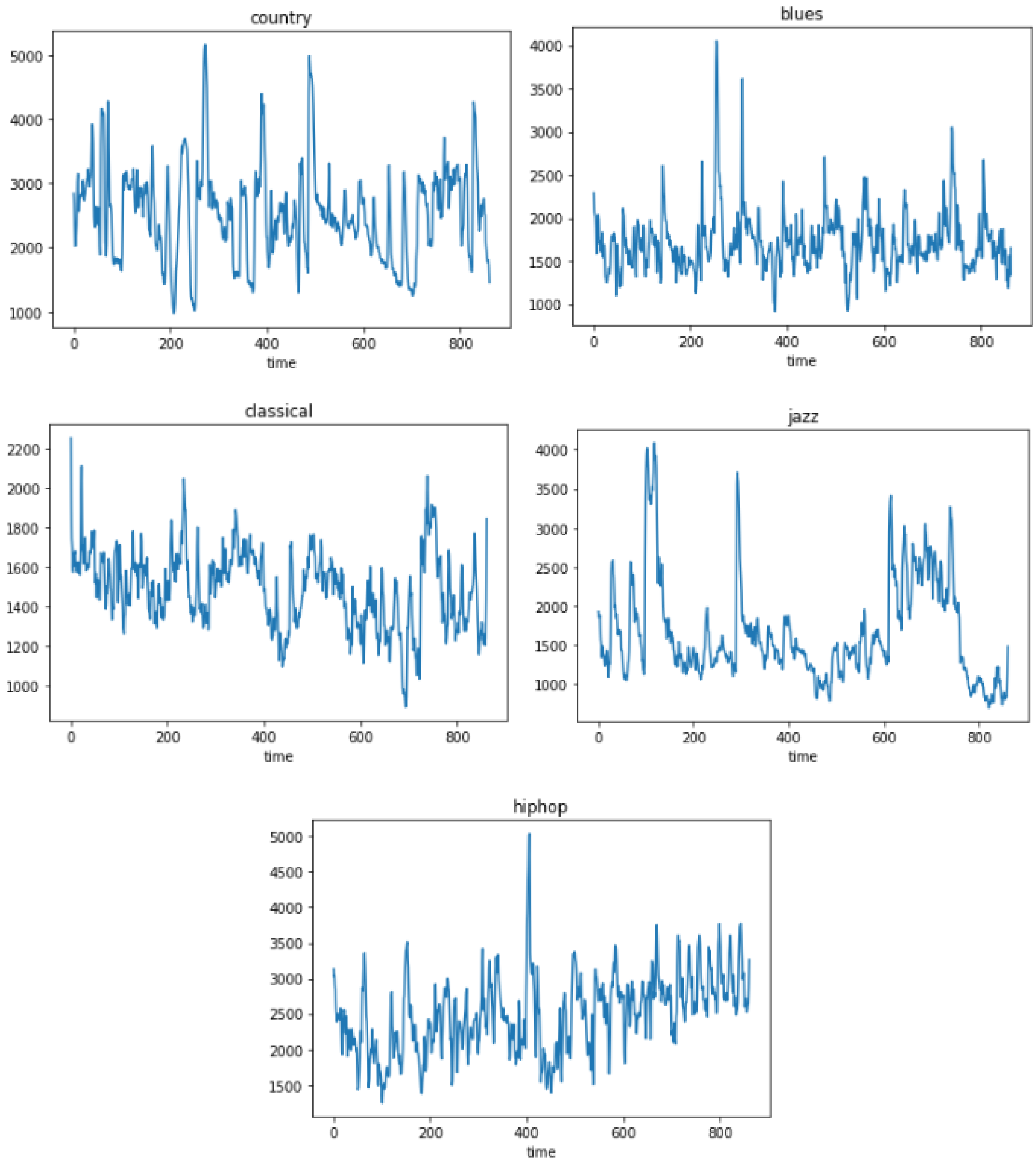

Fig 3: Spectral Centroid variation of the five genres

Spectral Roll-Off:

It is a measure of the shape of the signal. It represents the frequency below which a specified percentage of the total spectral energy, e.g. 85%, lies. This can help distinguish between a rock or metal, where the specified percentage of energy will correspond to a higher frequency to that of blues or classical where the spectral energy is concentrated at lower frequencies. Fig. 4 shows the Spectral Roll-off variation among the mentioned genres. Its X axis denotes the time and the Y axis denotes the spectral roll off point.

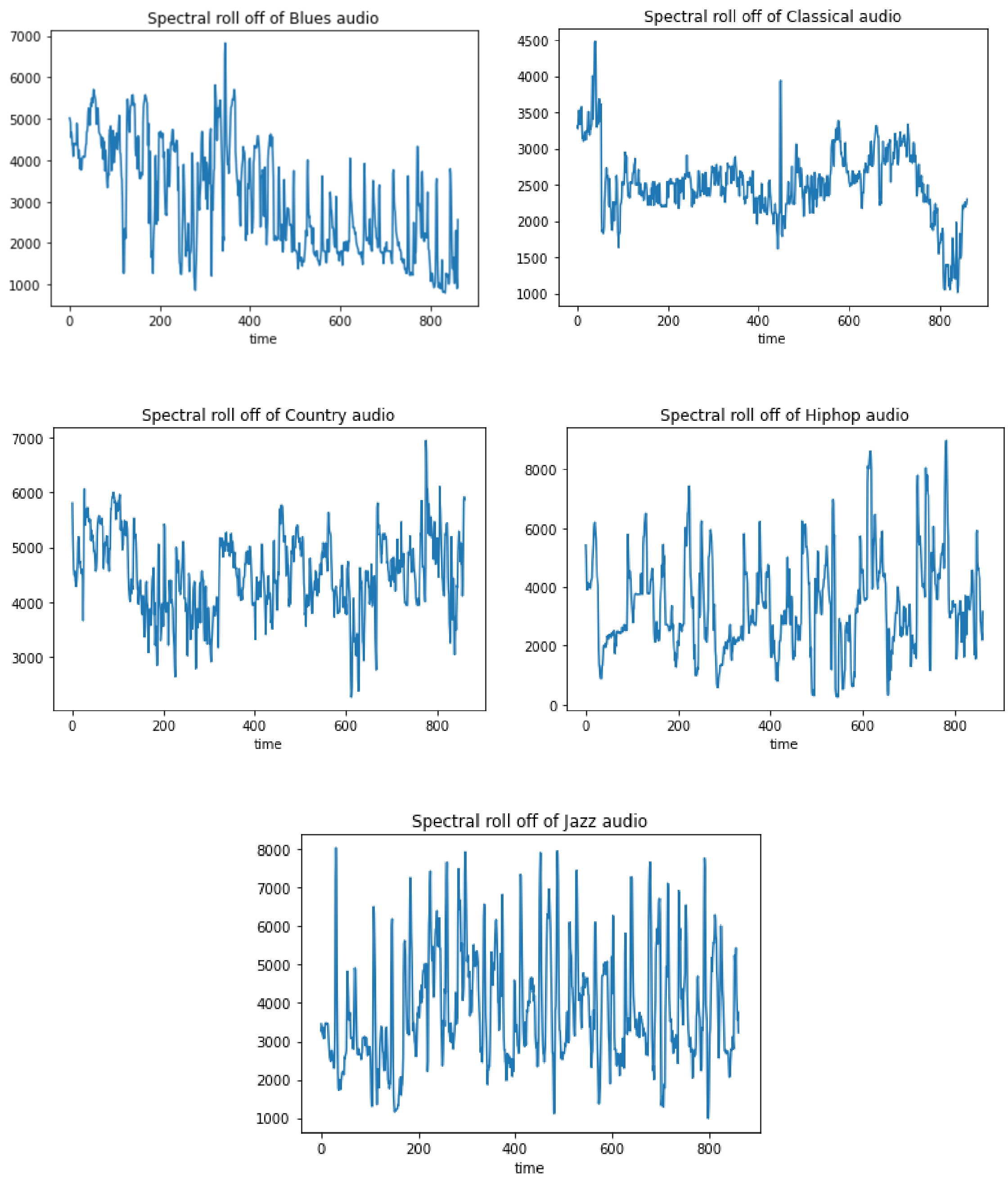

Fig 4:Spectral Roll off variation of the five genres

Mel Frequency Cepstral Coefficients:

The Mel frequency cepstral coefficients (MFCCs) of a signal are a small set of features (usually about 10–20) which concisely describe the overall shape of a spectral envelope. It models the characteristics of the human voice. The MFCCs encode the timbral properties of the music signal by encoding the rough shape of the log-power spectrum on the Mel frequency scale. Fig. 5 depicts the variation among the MFCCs of the mentioned genres. Its X axis denotes the MFCC coefficients and the Y axis denotes the MFCC values.

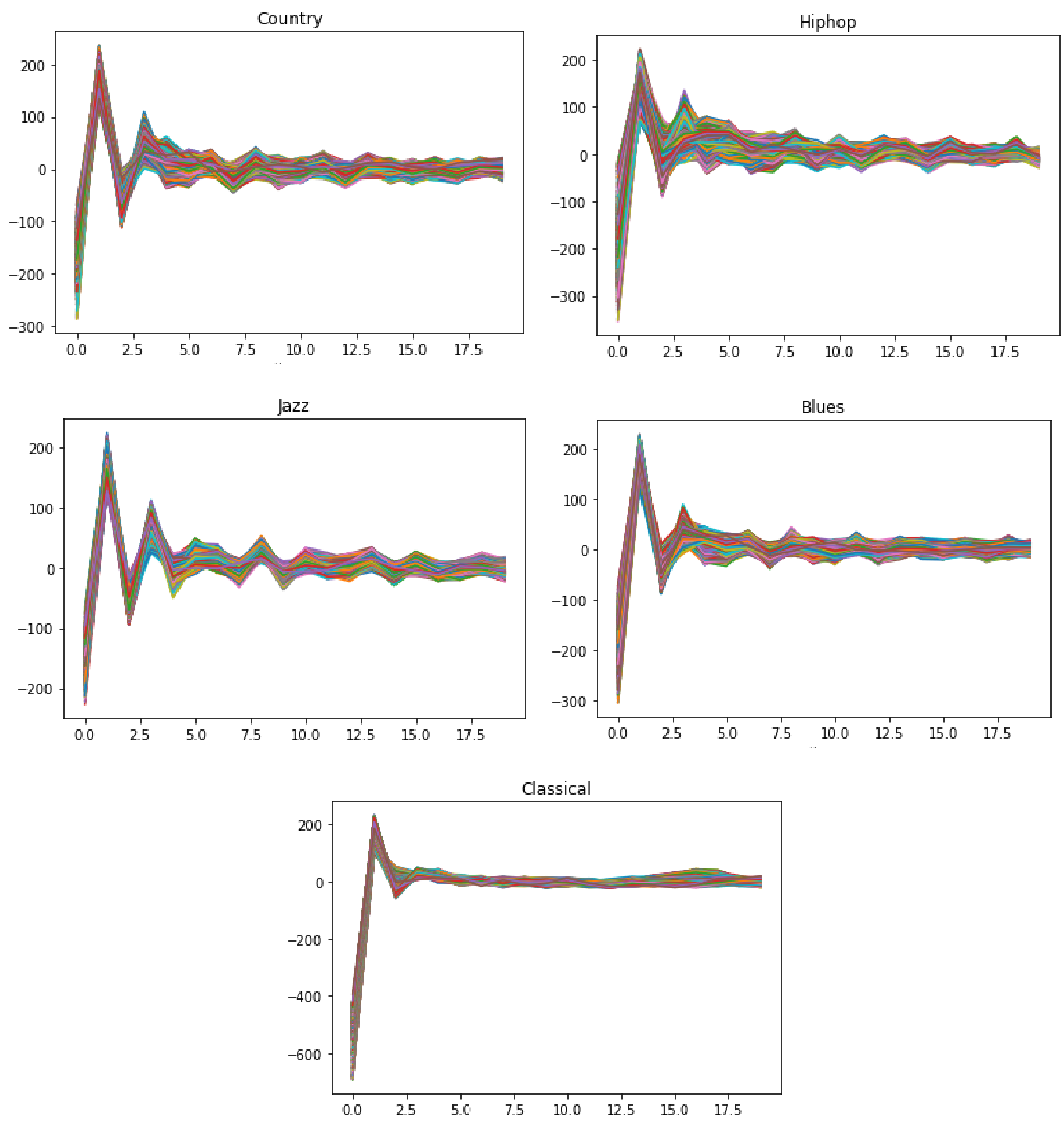

Fig 5: MFCC variation of the five genres

Chroma Shift:

Chroma features are an interesting and powerful representation for music audio in which the entire spectrum is projected onto 12 bins representing the 12 distinct semitones (or chroma) of the musical octave. One main property of chroma features is that they capture harmonic and melodic characteristics of music, while being robust to changes in timbre and instrumentation. Fig.6 shows the Chroma deviation variation among the mentioned genres. The X axis denotes the chroma coefficients and the Y axis denotes the chroma values.

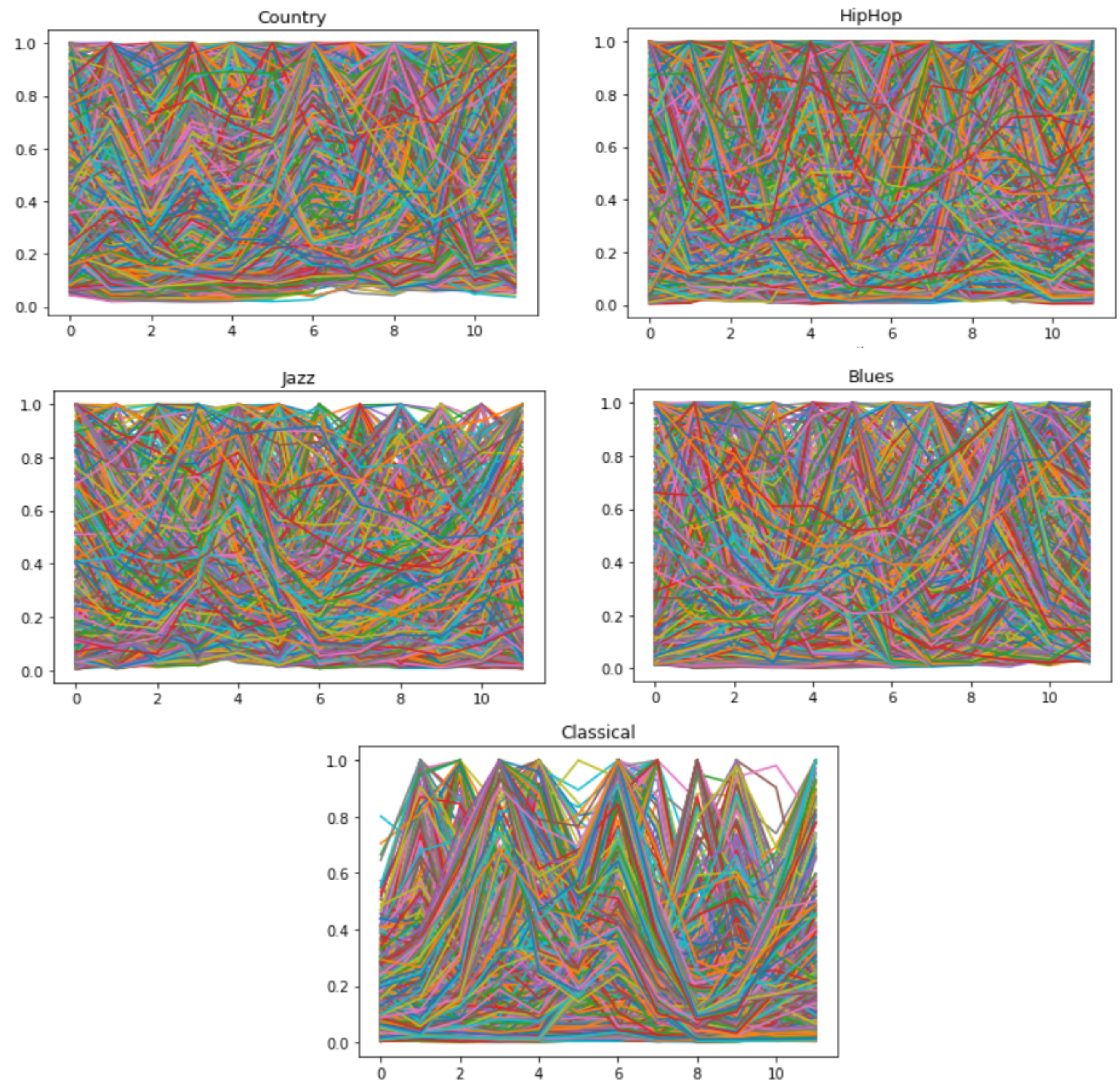

Fig 6: Chroma shift variation of the five genres

Other Features: Some other features like Spectral Bandwidth, RMSE, Harmony, Percussive and Tempo were also considered for the classification purposes.

3.4 Data Processing:

The raw data extracted was required to be processed before it could be used for the training purposes. This was done in order to correct for distortions and errors like radiometric errors in the data. A lot of the features were scaled to make the data on the same scale. Methods like the Standard scaler and min-max scaler from the Scikit-learn library were employed for feature scaling. We ensured that all the audio files were in 16 bit mono format using Audacity.

4. Classification Models:

In machine learning, classification refers to a predictive modelling problem where a class label is predicted for a given example of input data. A model will use the training dataset and will calculate how to best map examples of input data to specific class labels. As such the training dataset must be sufficiently representative of the problem and have many examples of each class label. The models used for the classification are described below.

4.1 K-nearest Neighbours:

The k-nearest neighbours (KNN) algorithm is a simple, easy-to-implement supervised machine learning algorithm that can be used to solve classification problems. In the k-nearest neighbours algorithm an object is classified by a vote of its neighbours, with the object being assigned to the class which is most common among its k nearest neighbours. If k = 1, then the object's class is that of the single nearest neighbour. Fig. 7 represents the K nearest neighbour model. Fig. 8 shows the classification using KNN model.

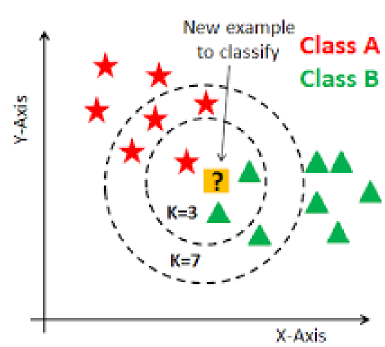

Fig. 7: Diagram of knn classification

```
In [14]: knn = KNeighborsClassifier(n_neighbors=5)

In [15]: knn.fit(X_train,Y_train)

Out[15]: KNeighborsClassifier()

In [16]: pred = knn.predict(X_test)
         print(pred)

[2 4 4 4 2 4 2 2 1 0 2 3 0 4 0 0 4 3 1 0 1 1 0 1 1 1 4 4 4 2 0 1 0 0 3 2 1
 1 1 0 1 2 3 4 1 1 3 1 1 0 3 3 1 1 1 2 0 2 2 4 3 3 0 0 2 1 1 4 4 4 3 3 2 1
 3 1 1 2 2 1 3 0 0 4 0 0 3 4 2 2 2 1 3 3 3 3 1 0 1 3 0 2 2 2 1 0 0 0 0 2 1
 4 2 1 0 1 2 1 0 3 2 1 4 2 0 3 2 2 2 0 0 0 1 3 1 0 3 4 0 0 1 0 4 3 0 1 2 4
 2 4 3 3 4 1 2 4 1 2 0 1 3 0 1 0 0 4 4 2 1 2 3 1 3 0 2 1 3 1 2 0 3 0 0 2 2
 3 4 1 1 0 4 4 4 0 0 4 0 0 3 2 3 1 4 3 3 0 3 1 1 0 2 3 3 2 2 1 2 1 0 0 2 3
 3 0 4 1 4 4 4 1 1 3 1 3 2 4 4 0 2 2 2 3 0 2 2 0 4 0 3 3 2 2 3 1 0 3 3 3 3
 3 0 4 3 4 3 1 0 1 3 4 0 1 2 3 2 3 1 3 3 2 4 2 4 0 3 0 4 0 4 2 4 4 4 4 3 1
 0 3 4 1]
```

Fig. 8: Knn model using scikit-learn

4.2 Logistic Regression:

Logistic regression is named for the function used at the core of the method, the logistic function.In order to map predicted values to probabilities, we use the sigmoid function. The function maps any real value into another value between 0 and 1. In machine learning, we use sigmoids to map predictions to probabilities. The sigmoid curve can be represented with the help of the following graph. We can see the values of the y-axis lie between 0 and 1 and cross the axis at 0.5. Fig. 9 shows the sigmoid function used.

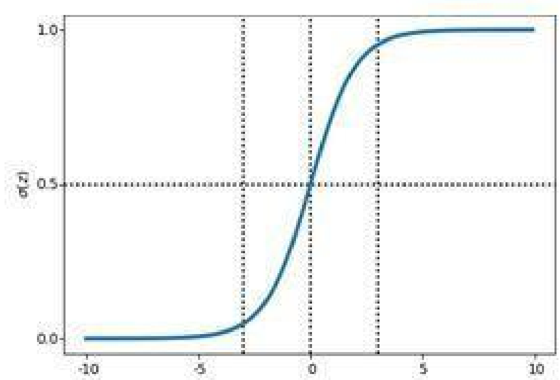

Fig. 9: Diagram of sigmoid function $f(x) = 1/(1 + e^{-x})$

We then calculate the likelihood of parameters. Likelihood is the probability of data, given a model and specific parameter values i.e. it measures the support provided by the data for each possible value of the regression coefficients. Finally we calculate the cost function and apply gradient descent algorithm to update the regression coefficients to achieve minimum cost value. Fig. 10 represents OneVOne and OneVRest Classification. Fig. 11 shows the code snippet for Logistic Regression.

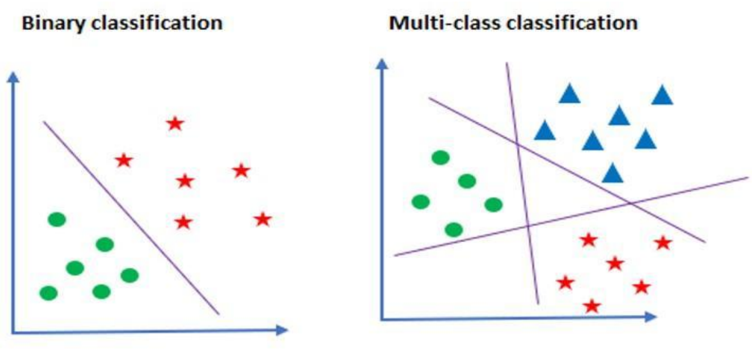

Fig 10: Diagram representing OneVOne and OneVRest Classification

```
In [35]: #model-logistic-regression
         from sklearn.linear_model import LogisticRegression
```

```
In [40]: logreg = LogisticRegression(C=0.01, solver='lbfgs', verbose=0 ,multi_class='auto').fit(x_train,y_train)
         logreg

Out[40]: LogisticRegression(C=0.01, class_weight=None, dual=False, fit_intercept=True,
                            intercept_scaling=1, l1_ratio=None, max_iter=100,
                            multi_class='auto', n_jobs=None, penalty='l2',
                            random_state=None, solver='lbfgs', tol=0.0001, verbose=0,
                            warm_start=False)
```

Fig 11: LR model using scikit-learn

4.3 Random Forest:

The main idea behind random forest is that it is a collection of a large number of decision trees that operate together to predict a class. Each individual decision tree in the random forest predicts a class and the class with the most number of votes becomes the predicted output of the random forest model. In our random forest model using scikit-learn the total number of trees in the forest i.e. estimators = 400, the splitting evaluation method i.e. criterion is chosen as 'gini' and max_depth is set to 20. Fig. 12 shows how Random Forest works. Fig. 13 has the code snippet for implementing Random Forest.

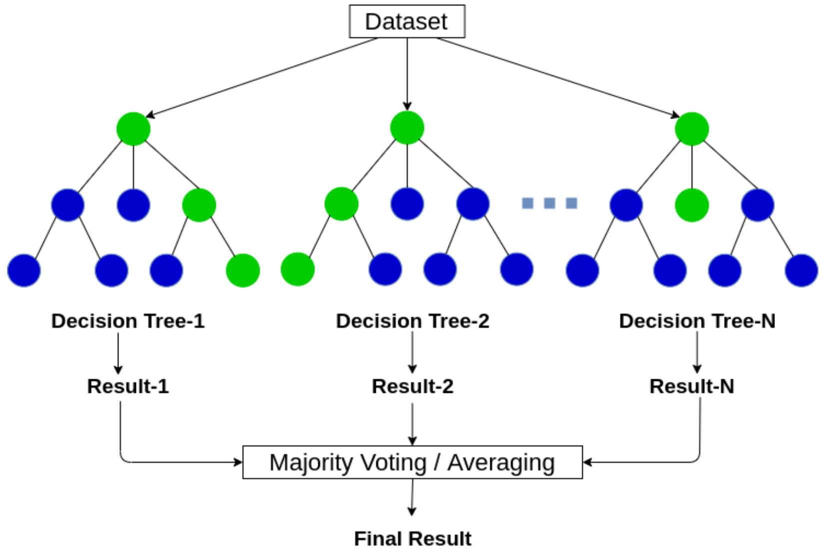

Fig 12: Diagram demonstrating working principle of random forest

```
In [8]: from sklearn import datasets
        from sklearn.metrics import accuracy_score
        from sklearn.ensemble import RandomForestClassifier

In [9]: model = RandomForestClassifier(n_estimators=400, max_depth=20)
        model.fit(x_train, y_train)

Out[9]: RandomForestClassifier(max_depth=20, n_estimators=400)
```

Fig 13: Random Forest model using scikit-learn

4.4 Artificial Neural Network:

An artificial neural network is like a human brain which consists of several nodes or neurons interconnected like a web. An artificial neural network consists of an input layer, an output layer and one or more hidden layers. The input layer receives several input data of different shapes and forms, the hidden layers extract various information from the input data and the model learns about the information presented and finally the output layer predicts the output. During the supervised phase the Ann model computes the error of the predicted values from the actual values. Then it uses backpropagation to update the weights and biases of the connections between the processing units until the error value ofis minimum. Each neuron or processing unit takes an input value, then processes the data using weights and biases and finally passes it through an activation function and converts the input into a more useful output. Fig. 14 shows how a neuron or processing unit works. Fig.15 shows what an artificial neural network looks like. Fig. 16 represents our artificial neural network model created using Keras.

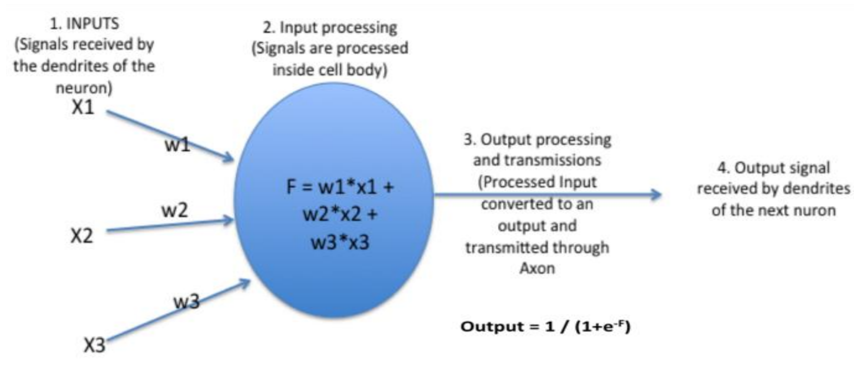

Fig 14: Diagram of a single processing unit

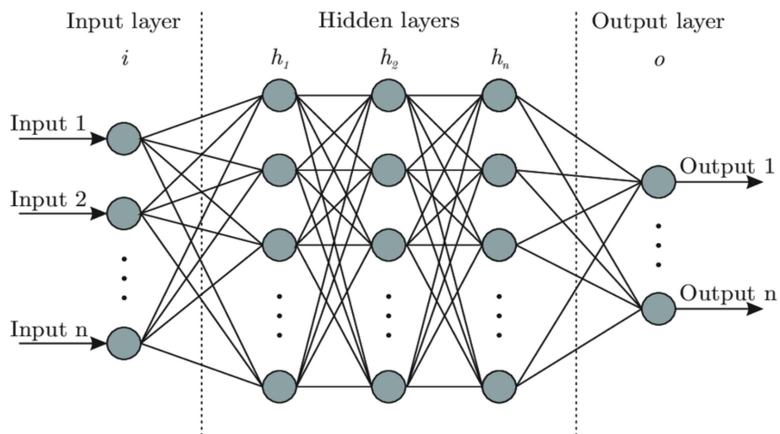

Fig 15: Diagram of an artificial neural network

```
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_40 (Dense)            (None, 256)               12032

 batch_normalization_32 (Bat  (None, 256)              1024
 chNormalization)

 dropout_32 (Dropout)        (None, 256)               0

 dense_41 (Dense)            (None, 128)               32896

 batch_normalization_33 (Bat  (None, 128)              512
 chNormalization)

 dropout_33 (Dropout)        (None, 128)               0

 dense_42 (Dense)            (None, 64)                8256

 batch_normalization_34 (Bat  (None, 64)               256
 chNormalization)

 dropout_34 (Dropout)        (None, 64)                0

 dense_43 (Dense)            (None, 32)                2080

 batch_normalization_35 (Bat  (None, 32)               128
 chNormalization)

 dropout_35 (Dropout)        (None, 32)                0

 dense_44 (Dense)            (None, 5)                 165

=================================================================
Total params: 57,349
Trainable params: 56,389
Non-trainable params: 960
```

Fig 16: Keras regression model.

5. Testing:

The GTZAN dataset has 100 audio files of each genre. We have calculated the training and validation accuracy of each model for five genres i.e. blues, classical, jazz, hip hop and country. We then selected the best model based on validation accuracy and then tested our model with unknown data i.e. audio files external to GTZAN dataset. The following table shows a comparative study of accuracy achieved by different classifiers.

Models	Segment	Validation Accuracy
K Nearest Neighbours	30sec	87%
Logistic Regression	30sec	86%
Random Forest	30sec	89%
Artificial Neural Network using Keras Regression	30sec	92.44%

Table 1: Comparative analysis of accuracy achieved by different classifiers.

From the above table it is clear that keras regression and random forest models are giving the best results.

Fig. 17 represents the accuracy for Keras regression model and Fig. 18 represents the confusion matrix.

```
model = load_model('Keras_reg_30sec_new5_35_AudioSegment.h5')
score = model.evaluate(X_test,y_test, verbose=0)
print("%s: %.2f%%" % (model.metrics_names[1], score[1]*100))
```

```
accuracy: 92.44%
```

Fig 17: Keras regression model accuracy (segment=30sec)

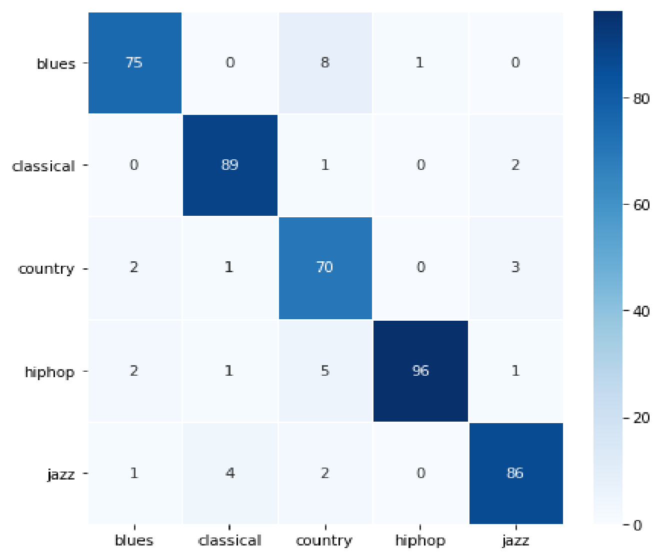

Fig 18: Confusion Matrix for Keras model output (segment=30sec)

Fig. 19 represents the accuracy for the Random Forest model and Fig. 20 represents the confusion matrix.

```
print(classification_report(y_test,y_predicted))
```

	precision	recall	f1-score	support
0	0.89	0.91	0.90	56
1	0.92	0.98	0.95	59
2	0.82	0.87	0.85	54
3	0.94	0.86	0.90	71
4	0.86	0.83	0.85	60
accuracy			0.89	300
macro avg	0.89	0.89	0.89	300
weighted avg	0.89	0.89	0.89	300

Fig 19: Random Forest model accuracy (segment=30sec)

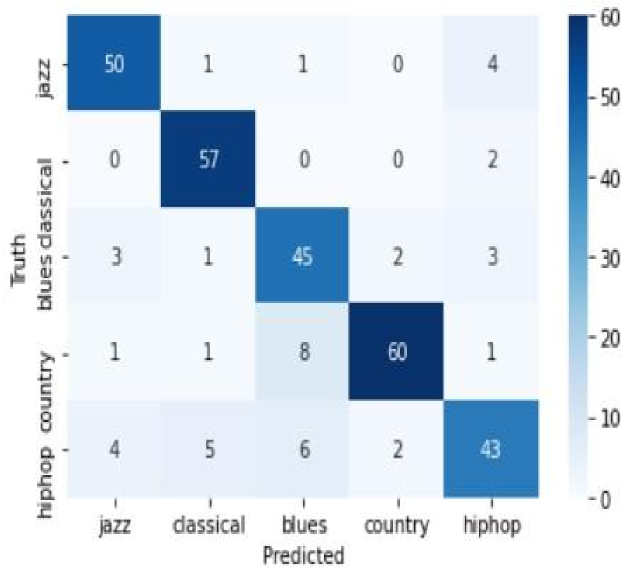

Fig 20: Confusion Matrix for Random Forest model output (segment=30sec)

6. Future Planning:

6.1 Scope for future improvement

The classification has been done with five genres. So we need to work on building a model that can give a good accuracy while working with all ten genres. Also we need to see how we can use LSTM to improve our accuracy. We can work on a model in the future that works as an ensemble consisting of multiple classifiers and the most voted prediction will be shown as the output. We also can create a web application which will process an uploaded audio file and predict the output genre. We can also include another feature in the web application that can create a playlist based on genre.

6.2 Conclusion

The project is an application of machine learning which is a trending technology. This project aims at categorising the music based on genres and finds various applications. We worked with only five genres and used classification techniques like KNN, Logistic Regression, Random Forest, Artificial Neural Network using Keras Regression. We are satisfied with the accuracy achieved by different classification models.

Gantt Chart for Project Planning:

Fig 19: Gantt Chart

References

1. Matthew Creme, Charles Burlin, Raphael Lenain, “Music Genre Classification ”, Stanford University, December 15, 2016.
2. Vishnupriya S, and K.Meenakshi, “Automatic Music Genre Classification using Convolutional Neural Network”, IEEE Conference 2018.
3. Arsh Chowdhry, “Music Genre Classification using CNN”,May 7, 2021.
4. Gautam Chettiar ,Kalaivani S, “ Music Genre Classification Techniques” Conference paper, Vellore Institute of Technology, 11 November, 2021.
5. Parul Pandey, “Music Genre Classification with Python”, December 13, 2018.
6. Introduction to Machine Learning by Alex Smola and S.V.N. Vishwanathan.
7. Benjamin Kedem. "Spectral analysis and discrimination by zero-crossings". Proceedings of the IEEE, 74(11):1477-1493, 1986.
8. Sanghamitra Das and Suchibrota Dutta and Dr. Arijit Ghoshal. “Stratification of Indian Dance Forms through Audio Signals”.
9. Tao Li and Mitsunori Ogihara. “A comparative study on content-based music genre classification”.
10. Hareesh Bahuleyan. “Music Genre Classification using Machine Learning Techniques”
11. Maximo Cobos, Jens Ahrens, Konrad Kowalczyk and Archontis Politis. “EURASIP Journal on Audio,

Speech, and Music Processing".

12. R. Thiruvengatanadhan, "Music Genre Classification using SVM",International Research Journal of Engineering and Technology(IRJET)''.

13. Ahmet Elbir1 , Hilmi Bilal Çam2 , Mehmet Emre İyican2 , Berkay Öztürk2 , Nizamettin Aydın1 , "Music Genre Classification andRecommendation by Using Machine Learning Techniques".