# VISVESVARAYA TECHNOLOGICAL UNIVERSITY

**'JNANA SANGAMA' BELAGAVI-590 018, KARNATAKA**

**PROJECT REPORT**

**ON**

## "MUSIC GENERE CLASSIFICATION USING MACHINE LEARNING"

**SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE AWARD OF THE DEGREE,**

**BACHELOR OF ENGINEERING**
**IN**
**INFORMATION SCIENCE & ENGINEERING**

**Submitted By**


1. **Deepika B**
   **1CG20IS010**
2. **Namratha T L**
   **1CG20IS028**
3. **Pavan T S**
   **1CG20IS031**
4. **Soundarya H R**
   **1CG20IS041**

**Under the guidance of:**

**Dr. Thara DK** Ph.D.,
Prof &Head, Dept. Of ISE,
CIT, Gubbi, Tumakuru.

**HOD:**

**Dr. Thara D K**, **Ph.D**
Prof & Head, Dept. of ISE,
CIT, Gubbi, Tumakuru.

**Channabasaveshwara Institute of Technology**
**(NAAC Accredited & ISO 9001:2015 Certified Institution)**
NH 206 (B.H. Road), Gubbi, Tumakuru – 572 216. Karnataka.


**(Affiliated to Visvesvaraya Technological University, Belagavi & Recognized by AICTE New Delhi)**

**2023-24**

# Channabasaveshwara Institute of Technology

(NAAC Accredited & ISO 9001:2015 Certified Institution)

NH 206 (B.H. Road), Gubbi, Tumakuru – 572 216. Karnataka.

**(Affiliated to Visvesvaraya Technological University, Belagavi & Recognized by AICTE New Delhi)**

**2023-24**

## DEPARTMENT OF INFORMATION SCIENCE & ENGINEERING

## CERTIFICATE

This is to certify that the project work entitled **"Music Genere Classification using machine learning"** has been successfully carried out by **Deepika B [1CG20IS010], Namratha T L [1CG20IS028], Pavan T S [1CG20IS031], Soundarya H R [1CG20IS041],** bonafide students of **CHANNABASAVESHWARA INSTITUTE OF TECHNOLOGY, GUBBI, TUMAKURU,**under our supervision and guidance and submitted in partial fulfillment of the requirements for the award of Degree in **Bachelor of Engineering** by **Visvesvaraya Technological University, Belagavi** during the academic year of 2023–2024. It is certified that all corrections/suggestions indicated for internal assessment have been incorporated in the report deposited in the departmental library. The project report has been approved as it satisfies the academic requirements for the above said degree.

**Guide:**

**Dr. Thara D K**,**Ph.D**
Prof &Head, Dept. of ISE,
CIT, Gubbi, Tumakuru.

**H.O.D:**

**Dr. Thara D K**,**Ph.D**
Prof & Head, Dept. of ISE,
CIT, Gubbi,Tumakuru.

**Principal:**

**Dr. Suresh D S Ph.D**
Director & Principal
CIT, Gubbi, Tumakuru.

**Examiners:**

1.

2.

Channabasaveshwara Institute of Technology
(NAAC Accredited & ISO 9001:2015 Certified Institution)
NH 206 (B.H. Road), Gubbi, Tumakuru – 572 216. Karnataka.
(Affiliated to Visvesvaraya Technological University, Belagavi & Recognized by AICTE New Delhi)

2023-24

**DEPARTMENT OF INFORMATION SCIENCE & ENGINEERING**

# UNDERTAKING

We the students **Deepika B [1CG20IS010], Namratha T L [1CG20IS028], Pavan T S [1CG20IS031], Soundarya H R [1CG20IS041]** of **VIII semester B.E.Information Science and Engineering** of **CHANNABASAVESHWARA INSTITUTE OF TECHNOLOGY, GUBBI,TUMAKURU** declare that Project work entitled**"MUSIC GENERE CLASSIFICATION USING MACHINE LEARNING"** has been carried out and submitted in partial fulfillment of the requirements for the award of degree in Bachelor of Engineering in **Information Science and Engineering** by the Visvesvaraya Technological University during the academic year 2023-2024.

1. **Deepika B**
   **[1CG20IS010]**
2. **Namratha T L**
   **[1CG20IS028]**
3. **Pavan T S**
   **[1CG20IS031]**
4. **Soundarya H R**
   **[1CG20IS041]**

**Channabasaveshwara Institute of Technology**

**(NAAC Accredited & ISO 9001:2015 Certified Institution)**

NH 206 (B.H. Road), Gubbi, Tumakuru – 572 216. Karnataka.

Partnering in Academic Excellence

**(Affiliated to Visvesvaraya Technological University, Belagavi & Recognized by AICTE New Delhi)**

**2023-24**

**DEPARTMENT OF INFORMATION SCIENCE & ENGINEERING**

# BONAFIDE CERTIFICATE

This is to certify that the Project work entitled **"MUSIC GENERE CLASSIFICATION USING MACHINE LEARNING"** is a bonafide work of **Deepika B [1CG20IS010], Namratha T L [1CG20IS028], Pavan T S [1CG20IS031], Soundarya H R [1CG20IS041]** students of **VIII** semester **B.E Information Science and Engineering** carried out at **Channabasaveshwara Institute of Technology**, Gubbi, Tumakuru, in partial fulfillment of the requirements of the award of degree in B.E. in **Information Science and Engineering** of Visvesvaraya Technological University, Belagavi under my supervision and guidance. Certified that to the best of my knowledge the work reported here in does not from part of any other thesis on the basis of which degree or award was conferred on earlier occasion to this or any other candidates.

**Guide:**

**Dr. Thara D K,Ph.D**
**Prof &Head, Dept. of ISE,**
**CIT, Gubbi, Tumakuru.**

# ACKNOWLEDGEMENT

A great deal of time and lot of effort has gone into completing this project report and documenting it. The number of hours spent in getting through various books and other materials related to this topic chosen by us have reaffirmed its power and utility in doing this project.

Several special people have contributed significantly to this effort. First of all, we are grateful to our institution "**Channabasaveshwara Institute of Technology"**, Gubbi which provided us an opportunity in fulfilling our most cherished desire of reaching the goal.

We acknowledge and express our sincere thanks to the beloved Director and Principal **Dr. Suresh D S** for his many valuable suggestions and continued encouragement and support in the academic endeavors.

We wish to express our deep sense of gratitude to **Dr. Thara D K,** Head of the Department of Information Science and Engineering for all the guidance and who still remains a constant driving force and motivated through innovative ideas with tireless support and advice during the project to examine and helpful suggestions offered.

We wish to express our deep sense of gratitude to our guide **Dr.Thara D K,** Asst. Prof, Dept of ISE, CIT, for her meticulous attention to details, which has contributed immeasurably to the quality of the project report.

We would express our gratitude towards our parents and friends for their kind cooperation and encouragement which helped us in completion of this project.

Finally, we would like to thank all the teaching and non-teaching staff of Dept of ISE, for their cooperation.

Thanking everyone….

**Deepika B [1CG20IS010]**

**Namratha T L [1CG20IS028]**

**Pavan T S [1CG20IS031]**

**Soundarya H R [1CG20IS041]**

# ABSTRACT

The music industry has undergone significant changes in recent years, both in its conventional existence and in the form of music created. The popularity of the music industry has grown, and with it, the market for different music styles. Music not only brings individuals together, but also provides insight into various cultures. To deliver better recommendations and suggestions to people, music needs to be classified into genres to meet the categorical needs of consumers. However, classifying music into different genres is a challenging task in the area of music information retrieval (MIR), and there have been several attempts to classify music using various machine learning approaches. The primary objective of this project is to automatically classify audio files into their respective musical genres. To achieve this goal, we will compare the performance of two classes of models: Support Vector Machines (SVMs) and Convolutional Neural Networks (CNNs). SVMs and CNNs are two commonly used approaches for classification under machine learning and deep learning, respectively, and have shown promise in delivering effective regression and classification results. The CNN model we will use is trained end-to-end to predict the genre label of an audio signal using its spectrograms, spectral Roll off, chroma features, and zero-crossing rate. We will use a dataset of audio tracks with similar sizes and frequency ranges, specifically the GTZAN genre classification dataset. In summary, this project aims to classify audio files into different genres using two models, SVMs and CNNs, and comparing their performance. The chosen dataset is the GTZAN genre classification dataset, and the low-level features of frequency and time domain will be used to classify the audio files

# CONTENTS

# CHAPTER 1

# INTRODUCTION

The development in the field of technology has led to the widespread adoption of machine learning solutions across multiple industries. The importance of categorizing music by genre rests in its fundamental function in creating reliable recommendation systems for music organizations. Numerous machine and deep learning algorithms have been used throughout the years to conduct substantial study on the classification of musical genres. Machine learning algorithms have the potential to analyse large datasets automatically and identify interesting patterns. The field of genre classification and music information retrieval has 4 been the subject of extensive research, benefiting both the music industry and consumers. Digitalization has revolution the music industry, from the creation and production to the consumption and sharing of music. Music is categorized into subjective categories called genres, and human experts have traditionally been relied upon to attribute genre tags to songs. With the ever-growing customer base, there has been an increase in demand for various music styles. Music not only brings people together, but it also provides insight into different cultures. Hence, it is essential to categories music according to genres to satisfy the needs of people categorically. A substantial amount of research has been done on classifying music genres, which may be divided into two types. The GTZAN and FMA databases are two that are often used. Deep learning has been an effective method for creating reliable end-to-end systems in a variety of image, video, audio, and voice analysis domains. This study is subject to certain limitations that relate to the quality of the data or the sampling method used, as well as the restricted size of the dataset. Additionally, there may be challenges in terms of feature engineering or optimizing the selected feature set. It's also vital to keep in mind that not all methods for the task at hand may have been included in this study. The GTZAN music audio file collection, which is widely used by music information retrieval researchers, was used as the dataset for this study. Ten different music genres were used for classification, including Jazz, Rock, Pop, Country, Hip-Hop, Reggae, Classical, Rhythm and Blues, Metal, and Disco.

## 1.1 OBJECTIVES

- To automatically assign genre labels to music tracks without human intervention, enabling efficient organization and retrieval of music in large databases.
- To enhance music recommendation systems by accurately identifying the genre of a song and recommending similar tracks based on genre preferences.
- To create personalized playlists for users based on their preferred music genres, improving user experience and engagement with music streaming platform.
- To facilitate music discovery by enabling users to explore new genres based on their existing preferences and listening habits.
- To extract relevant features from audio signals such as timbre, rhythm, and spectral characteristics, and analyze them to discern patterns specific to different music genres.

## 1.2 PROBLEM STATEMENT

To produce a model based on machine learning that solves the challenge of automatic classification of songs into its corresponding genre.

## 1.3 SCOPE OF THE PROJECT

- This task has broad applications in music recommendation systems, music streaming platforms, and automatic playlist generation.
- It typically involves feature extraction from audio signals, such as spectrogram analysis or timbral features, followed by the application of machine learning algorithms, such as support vector machines, neural networks, or decision trees, to classify the music into genres.
- Challenges in this domain include dealing with the subjective nature of music genres, handling diverse musical styles and subgenres, and addressing the high dimensionality of audio data.
- Despite these challenges, advances in machine learning techniques, particularly deep learning, have led to significant progress in music genre classification, enabling more accurate and robust systems.

# CHAPTER 2

# LITERATURE SURVEY

1. Music Genre Classification and Recommendation by Using Machine Learning Techniques Authors: Ahmet Elbir ; Hilmi Bilal Çam ; Mehmet Emre Iyican ; Berkay Öztürk ; Nizamettin Aydin, 2018 Innovations in Intelligent Systems and Applications Conference (ASYU) Music genre prediction is one of the topics that digital music processing is interested in. study, acoustic features of music have been extracted by using digital signal processing techniques and then music genre classification and music recommendations have been made by using machine learning methods. In addition, convolutional neural networks, which are deep learning methods, were used for genre classification and music recommendation and performance comparison of the obtained results has been. In the study, GTZAN database has been used and the highest success was obtained with the SVM algorithm.

2. Music Genre Classification: A N-Gram Based Musicological Approach Authors: Melody Moh ; Teng-Sheng Moh, 2017 IEEE 7th International Advance Computing Conference (IACC) Digitalization of music has grown deep into people's daily life. Derived services of digital music, such as recommendation systems and similarity test, then become essential for online services and marketing essentials. As a building block of these systems, music genre classification is necessary to support all these services. Previously, researchers mostly focused on low-level features, few of them viewed this problem from a more interpretable way, i.e., a musicological approach. This creates the problem that intermediate stages of the classification process are hardly interpretable, not much of music professionals' domain knowledge was therefore useful in the process.

This paper approaches genre classification in a musicological way. The proposed method takes into consideration the high-level features that have clear musical meanings, so that music professionals would find the classification results interpretable. To examine more musicological elements other than additional statistical information, we use a dataset of only symbolic piano works, including more than 200 records of classical, jazz, and ragtime music. Feature extraction and n-gram text classification algorithm are performed. The proposed method proves its concept with experimental results achieving the prediction accuracy averaged above 90%, and with a peak of 98%.

3. Long short-term memory recurrent neural network-based segment features for music genre classification Authors: Jia Dai ; Shan Liang ; Wei Xue ; Chongjia Ni ; Wenju Liu, 2016 10th International Symposium on Chinese Spoken Language Processing (ISCSLP) In the conventional frame feature-based music genre classification methods, the audio data is represented by independent frames and the sequential nature of audio is totally ignored. If the sequential knowledge is well modeled and combined, the classification performance can be significantly improved. The long short-term memory (LSTM) recurrent neural network (RNN) which uses a set of special memory cells to model for long-range feature sequence, has been successfully used for many sequence labeling and sequence prediction tasks. In this paper, we propose the LSTM RNN based segment features for music genre classification. The LSTM RNN is used to learn the representation of LSTM frame feature. The segment features are the statistics of frame features in each segment. Furthermore, the LSTM segment feature is combined with the segment representation of initial frame feature to obtain the fusional segment feature. The evaluation on ISMIR database show that the LSTM segment feature performs better than the frame feature. Overall, the fusional segment feature achieves 89.71% classification accuracy, about 4.19% improvement over the baseline model using deep neural network (DNN). This significant improvement shows the effectiveness of the proposed segment feature.

4. Music Genre Recognition Using Residual Neural Networks Authors; Dipjyoti Bisharad ; Rabul Hussain Laskar TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON) Genre is an abstract, yet a characteristic feature of music. Existing works for automatic genre classification compute a set of features from the audio and design classifier on top of it. Such models, in general, compute these features over a relatively long duration of the audio. In this paper, a residual neural network-based model is proposed for genre classification which is trained on short clips of just 3 seconds duration. Also, traditional genre classification algorithms will assign a single genre to an audio clip. However, it well established that different genres have overlapping characteristics. Considering this ambiguous nature of the genre, the model proposed in this work can assign three genre labels to a music clip, with each genre associated with some probability. The proposed model has an error rate of 18%, 9%, and 5.5% while predicting into top-1, top-2 and top-3 genres for a music clip respectively.

5. On Combining Diverse Models for Lyrics-Based Music Genre Classification Authors: Caio Luiggy Riyoichi Sawada Ueno ; Diego Furtado Silva, 2019 8th Brazilian Conference on Intelligent Systems (BRACIS) Automatic music organization and retrieval is a highly required task nowadays. Labeling songs with summarized but descriptive information have implications in a wide range of tasks in this scenario. The genre isone of the most common labels used for music recordings. Using this piece of information, music platforms can organize collections by, for instance, associating songs and artists with similar characteristics. Lyrics represent an alternative source of data for genre recognition.

6. "A Comparative Evaluation of Feature Representations in Music Genre Classification" Authors: Juan P. Bello, Chris Duxbury, and Mark B. Sandler (2005): This study compared the effectiveness of various feature representations, including spectrogram-based features, cepstral coefficients, and rhythm histograms. Different classifiers such as k-NN, SVMs, and decision trees were evaluated on a diverse dataset spanning multiple genres. Results highlighted the importance of feature selection and preprocessing techniques in improving classification accuracy.

7. "Music Genre Classification: A Taxonomy and Evaluation" Authors: G. Tzanetakis and P. Cook (2002): This seminal work presented a taxonomy of music genres and evaluated the performance of different classification algorithms, including k-nearest neighbors, support vector machines, and neural networks. They used a dataset consisting of audio features such as spectral centroid, zero-crossing rate, and Mel frequency cepstral coefficients (MFCCs).Results showed that SVMs and k-NN classifiers performed well across various genres.

# CHAPTER 3

# SYSTEM ANALYSIS

## 3.1 EXISTING SYSTEM

The existing systems for music genre classification using machine learning typically involve extracting features from audio signals, such as Mel-frequency cepstral coefficients (MFCCs), spectral features, or rhythmic patterns, and then training machine learning models, like support vector machines (SVMs), random forests, or deep neural networks, to classify these features into predefined music genres. These models are trained on labeled datasets containing audio samples with their corresponding genres. The performance of these systems depends on the quality and size of the training dataset, the choice of features, and the complexity of the classification model.

## 3.2 PROPOSED SYSTEM

proposed system of music genre classification using machine learning would involve several key steps. Initially, a comprehensive dataset of audio samples spanning various genres would be collected and preprocessed to extract relevant features, such as spectral characteristics, tempo, and rhythm pattern. These features would then be used to train a machine learning model, employing algorithms like support vector machines, decision trees, or neural networks. The model would learn to distinguish between different music genres based on the extracted features. Following training, the system would be evaluated using a separate test dataset to assess its accuracy and generalization performance. Finally, the trained model could be deployed in real-world applications, such as music recommendation systems or automatic genre tagging for digital music libraries. Ongoing refinement and optimization would be essential to improve the system's performance and adaptability to diverse musical styles and characteristic

# CHAPTER 4

# SYSTEM DESIGN

## 4.1 SYSTEM ARCHITECTURE

The design of system deals with how system is developed. It explains the flow of functionalities in brief. The section contains system data flow diagram, flow chart and sequence diagrams described below.

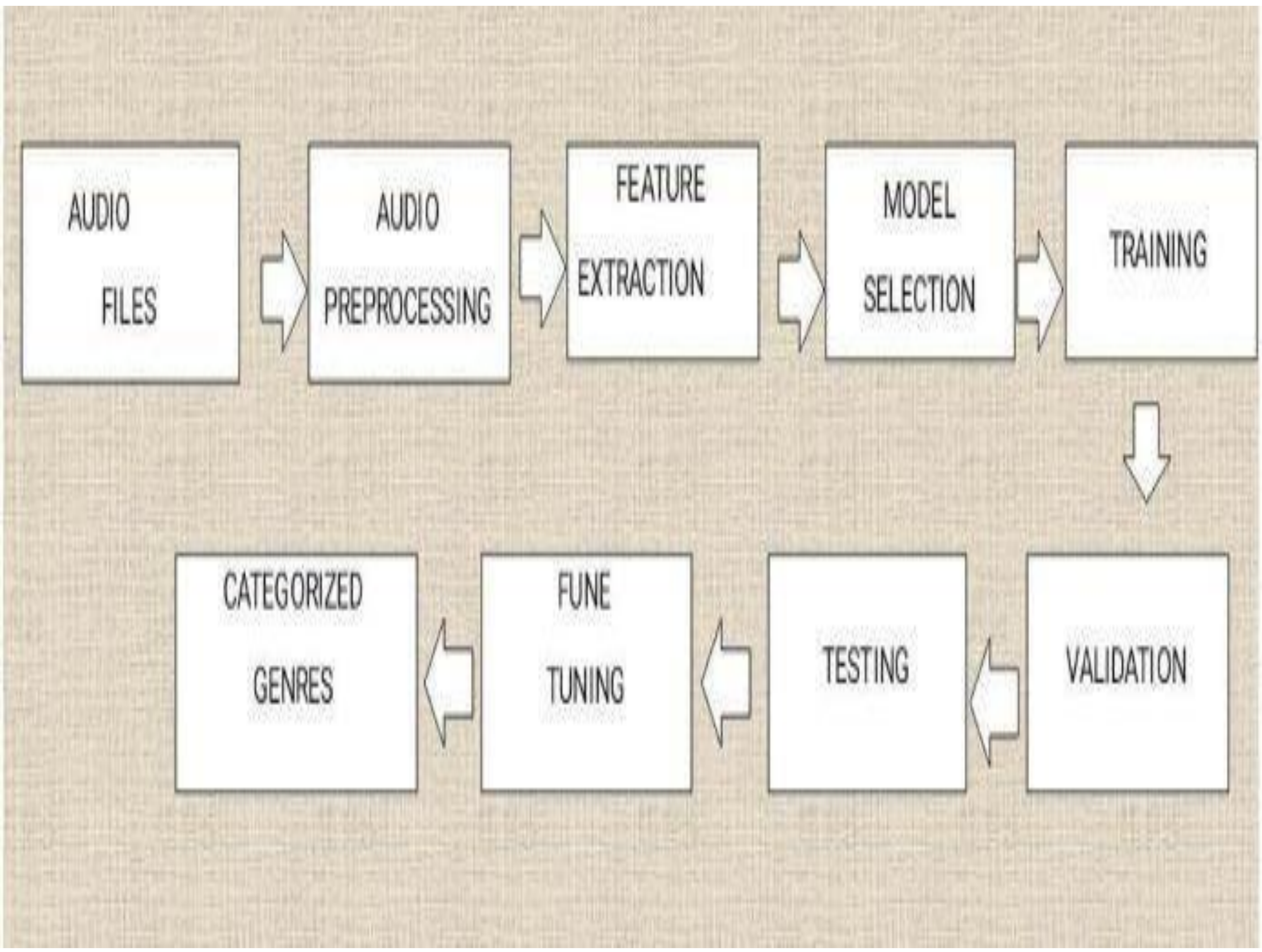



Fig 4.1. System Architecture

## 4.2 Flow Chart

A flow chart is a type of diagram that represents a workflow or process. A flowchart can also be defined as a diagrammatic representation of an algorithm, a step by step approach to solving a task. The flow chart shows the steps as boxes of various kinds, and their order by connecting the boxes with arrows.

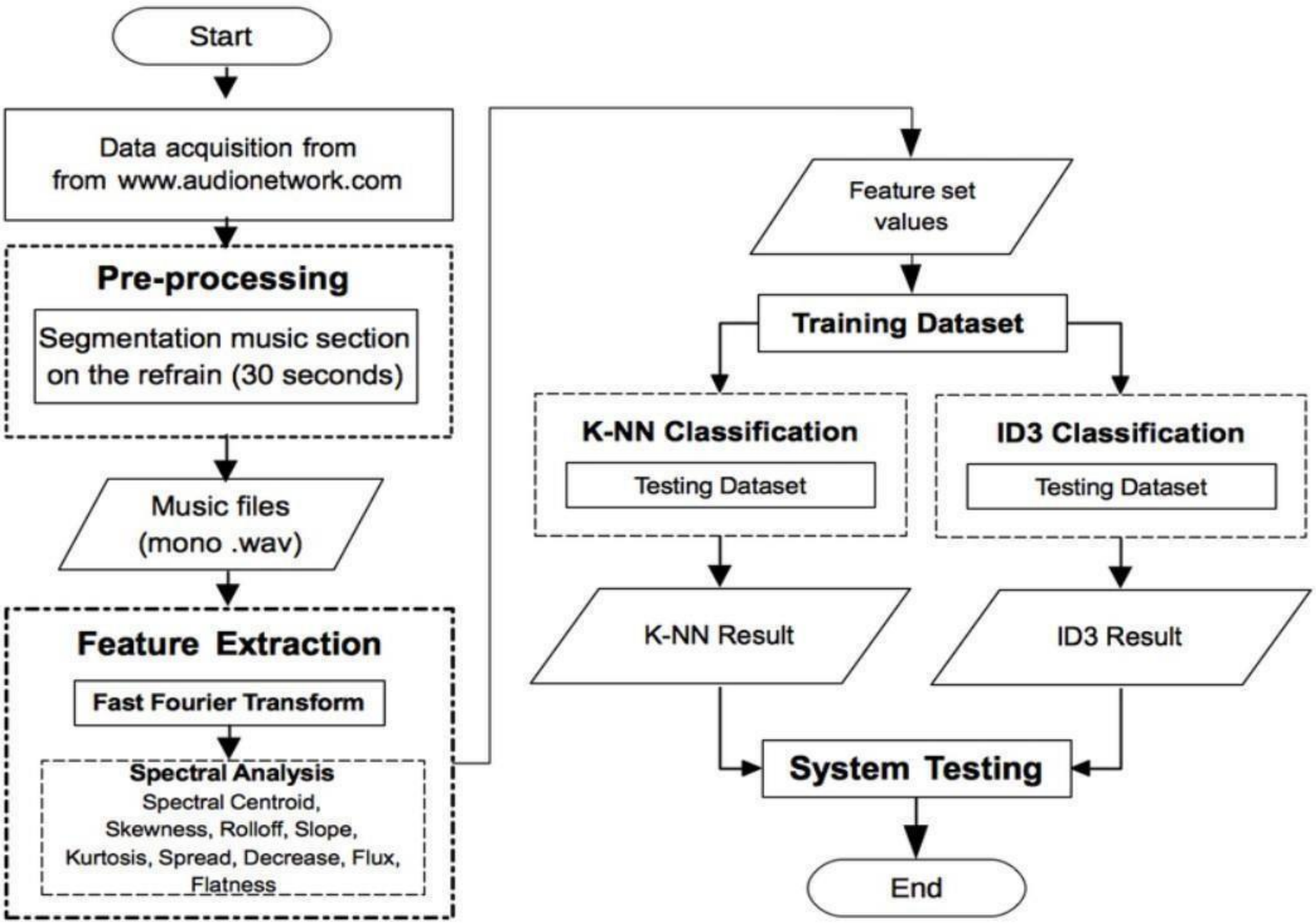



Fig 4.2. Flowchart of Fetal Abnormalities prediction

## 4.3 Sequence diagram

A sequence diagram is a type of interaction diagram because it describes how and in what order a group of objects works together. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario. Sequence diagrams are typically associated with use case realizations in the logical view of the system under development. Sequence diagrams are sometimes called event diagrams

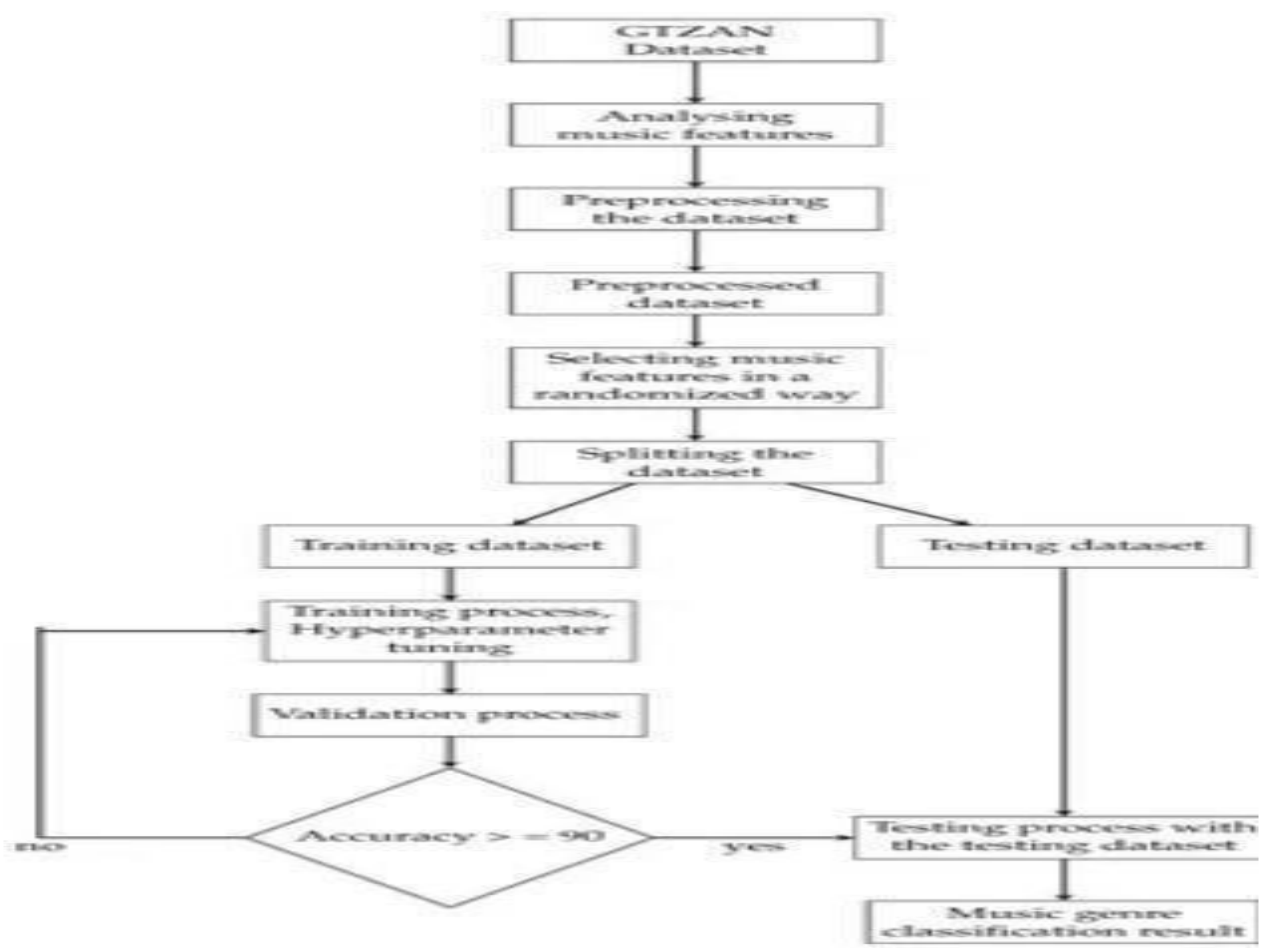



Fig 4.3.Sequence diagram

# CHAPTER 5

## IMPLEMENTATION

The proposed methodology comprises of following phases:

### 5.1 AUDIO SIGNALS

In the context of music genre classification, an audio file can be defined as a digital representation of sound, typically stored in formats like MP3, WAV, or FLAC, which contains the sonic characteristics and attributes used to identify and classify different genres of music.

### AUDIO PREPROCESSING

Audio preprocessing for music genre classification involves transforming raw audio data into a format that is suitable for analysis and classification. This typically includes steps such as converting audio into a spectrogram representation, applying techniques like windowing and normalization to enhance features, and possibly reducing dimensionality through methods like feature extraction or transformation to capture relevant characteristics of different music genres.

### Feature Extraction

Feature extraction in the context of music genre classification involves the process of transforming raw audio data into a set of meaningful numerical features that capture different aspects of the audio signal, suchas timbre, rhythm, and spectral content. These features serve as input to machine learning algorithms for training and classification tasks, enabling the identification and categorization of music into different genresbased on their distinctive acoustic characteristics.

Commonly used features for music genre classification includes;

a. Mel-frequency cepstral coefficients (MFCCs):
    i. These capture the spectral envelope of a sound and are widely used in speech and music processingtasks.
b. Spectral features:
    i. Including spectral centroid, spectral bandwidth, spectral contrast, and spectral roll-off, whichdescribe the distribution of frequencies in a signal.

c. Rhythm feature:
    i. Such as tempo, beat, and rhythm patterns, which capture the temporal structure of music.

d. Harmonic features:
    i. Including chroma features, which represent the energy of different pitch classes in a musicalsegment, and harmonic/percussive component separation.

e. Statistical features:
    i. Such as mean, variance, skewness, and kurtosis of various audio descriptors, which provideinformation about the statistical distribution of audio signals.

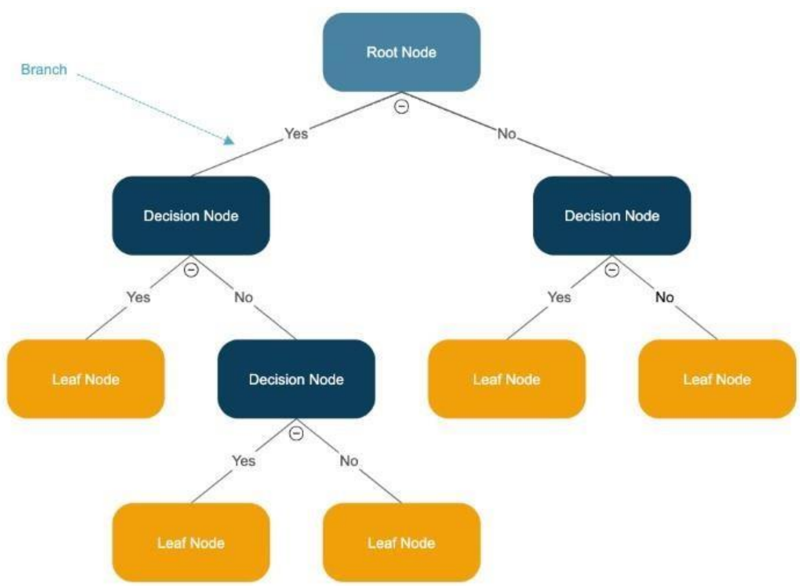
f. Temporal features:
    i. Such as zero-crossing rate and time- domain statistics, which capture characteristics of the audio signal over time.

**Model Selection:**

Model selection refers to the process of choosing the most suitable machine learning algorithm or model to accurately classify music into different genres. This involves evaluating various models based on their performance metrics, such as accuracy, precision, recall, and F1 score, to determine which one provides the results for the specific task of classifying music genres.

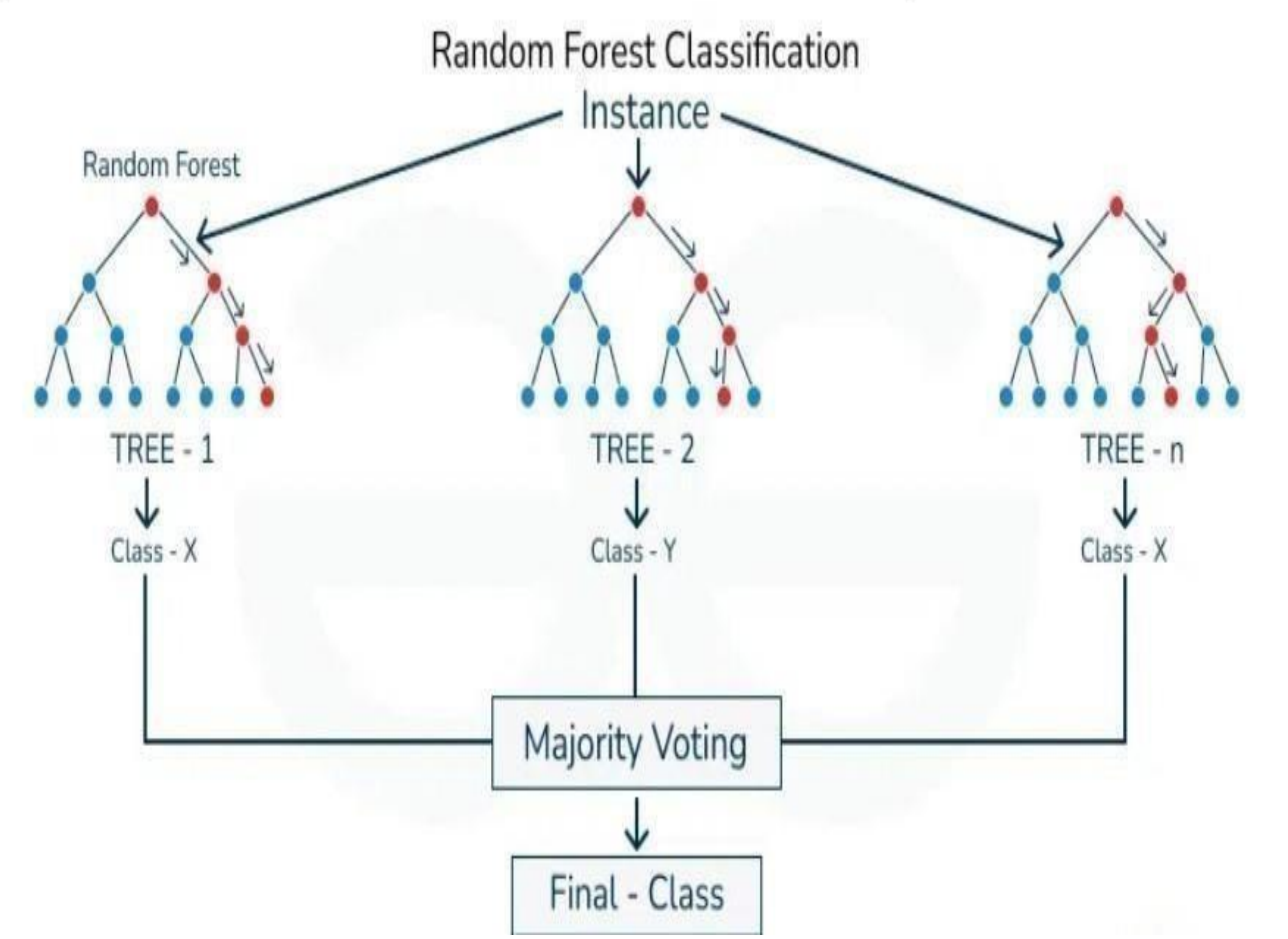
**Decision Tree:**

Involves structuring a tree-like graph where internal nodes represent features of the music (like tempo, key, etc.), and the branches represent the decision rules based on those features. At the leaf nodes, you have the predicted genre labels. You'd train the model with a dataset containing labeled music samples, then use it to classify new music samples into genres based on their features. It's a straightforward approach, but its performance depends on the quality of the features and the complexity of the decision rules.



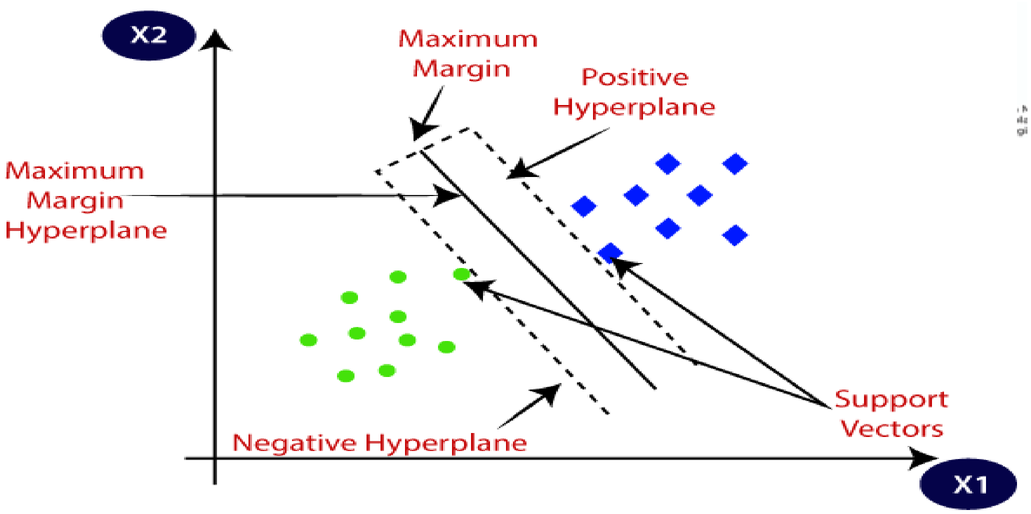
**Random Forest:**

Random forest is a popular machine learning algorithm used in music genre classification tasks. It works well because it can handle high-dimensional data (like audio features) and tends to generalize well to unseen data. With its ensemble of decision trees, it can capture complex relationships between features and labels, making it effective for classifying music into different genres.



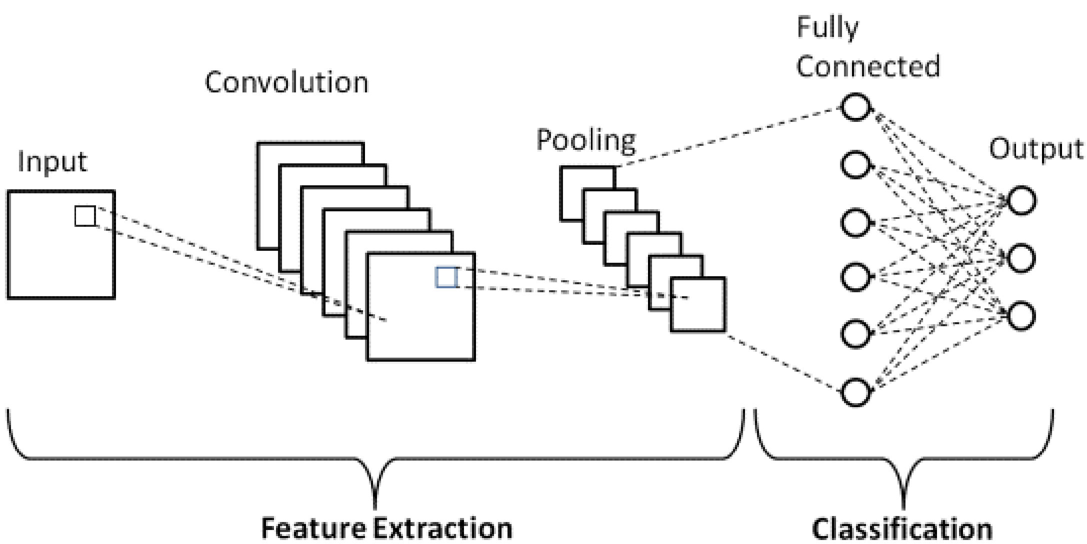

**SVM (Support vector Machine):**

Support vector machines (SVMs) are often used in music genre classification due to their ability to handle high- dimensional data effectively. In this context, features extracted from audio signals, such as spectral characteristics, tempo, and rhythm, are used to train the SVM model. SVMs excel in finding optimal decision boundaries between different classes, making them suitable for tasks like music genre classification where the classes may not be linearly separable. By learning from labeled examples, SVMs can classify new songs into predefined genres with high accuracy.



**CNN (Convolutional Neural Network):**

A CNN (Convolutional Neural Network) model in music genre classification works by analyzing spectrograms or other representations of audio data to extract features like timbre, rhythm, and pitch. These features are then processed through convolutional layers, which can detect patterns in different parts of the spectrogram. As the network goes deeper, it learns increasingly complex representations of the audio, eventuallyoutputting probabilities for each genre. Training involves feeding the network labeled examples of audio data, adjusting the model's parameters to minimize classification errors.

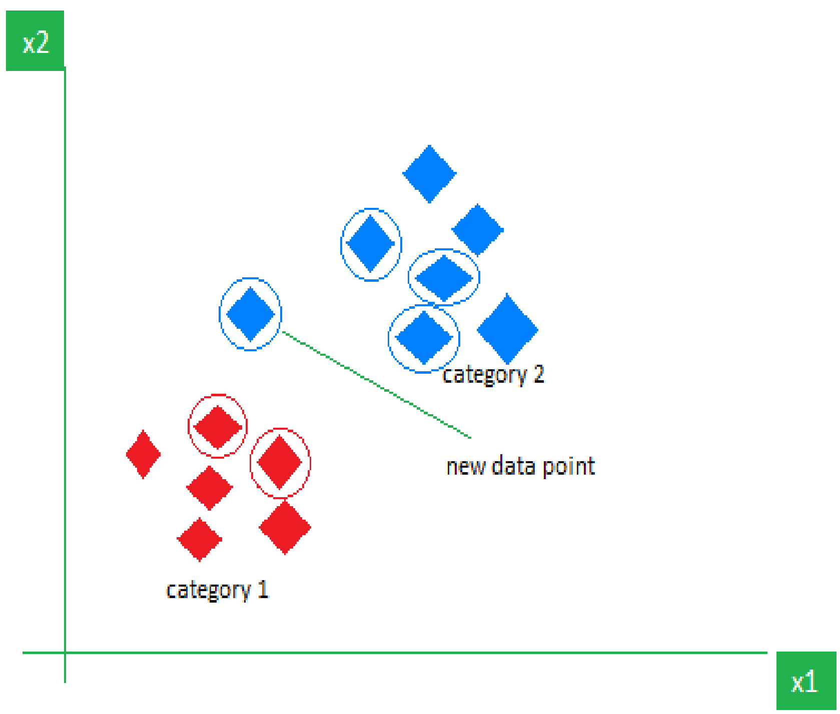



**Training**

The process of feeding a machine learning model with data (such as audio samples or features extracted from them) along with their corresponding genre labels, in order to teach the model to recognize patterns and characteristics specific to each genre. This enables the model to make predictions or classifications on new, unseen music samples based on the learned patterns.

**KNN (k-nearest neighbor):**

K-nearest neighbors (KNN) is a popular algorithm for music genre classification. In KNN, genres are assigned to songs based on the genres of their nearest neighbors in the feature space. Features like rhythm, melody, and timbre are commonly used to represent songs. KNN is simple to implement and understand, making it a good starting point for music classification tasks.



**Naive Bayes:**

Naive Bayes is a popular algorithm used in music generic classification. It works by calculating the probability that a given piece of music belongs to a certain genre based on the occurrence of certain features, such as tempo, pitch, and rhythm. It's "naive" because it assumes that all features are independent of each other, which may not always be the case in real world scenarios. However, despite its simplicity, **Naïve** Bayes can perform surprisingly well in music genre classification tasks, especially when combined with other techniques or algorithms.

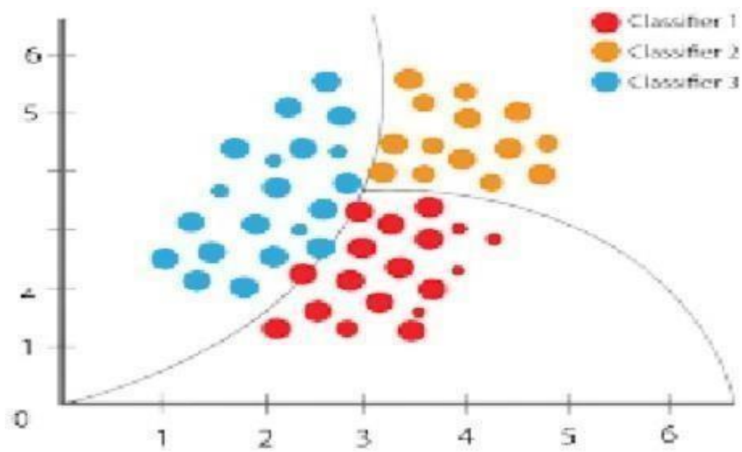



Fig 5.1 KNN Algorithm

**Validation**:

Validation refers to the process of assessing the accuracy and reliability of a classification model's predictionson music genre labels. This involves techniques such as cross-validation, where the model's performance is evaluated using different subsets of the dataset, and metrics like precision, recall, and F1-score are used to quantify its effectiveness in classifying music into specific genres.

**Testing:**

The testing process involves evaluating algorithms or models designed to automatically categorize music into predefined genres. It includes tasks like assessing the accuracy, precision, recall, and F1-score of the classification results, as well as considering factors like feature selection, model complexity, and dataset bias specific to the nuances of music genres.

**Fune tuning:**

Fine-tuning in music refers to the process of adjusting specific parameters within a musical composition or performance to align with the stylistic conventions and characteristics of a particular genre, enhancing its authenticity.

**Categorized Genres:**

Categorizing music based on various characteristics such as rhythmic, melodic, harmonic, and timbral elements, as well as cultural and historical contexts. It aims to provide a systematic framework for organizing and describing different styles of music, allowing for easier communication, analysis, and interpretation withinthe musical community

# CHAPTER 6

# RESULTS

## SCREENSHOTS:

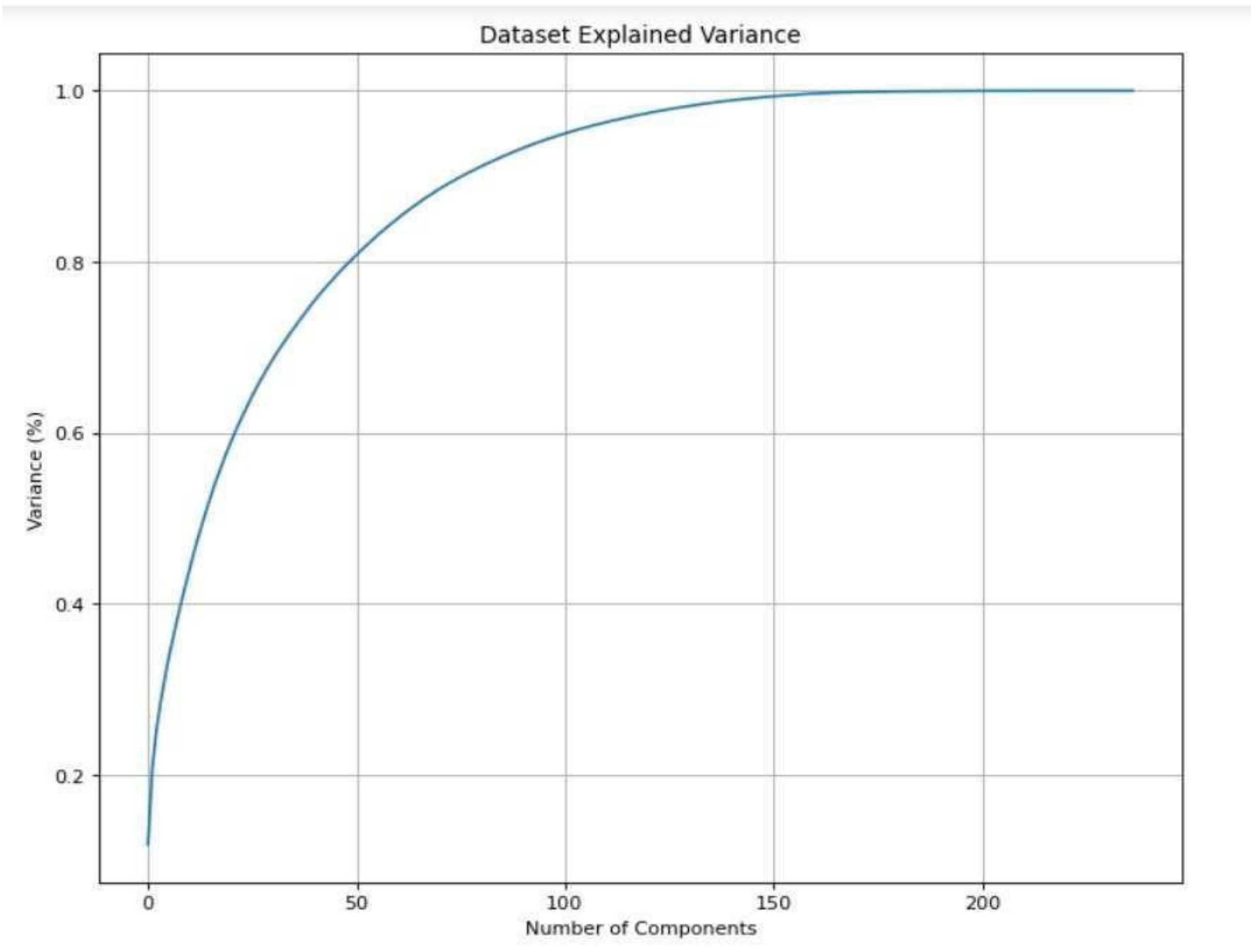



Fig 6.1

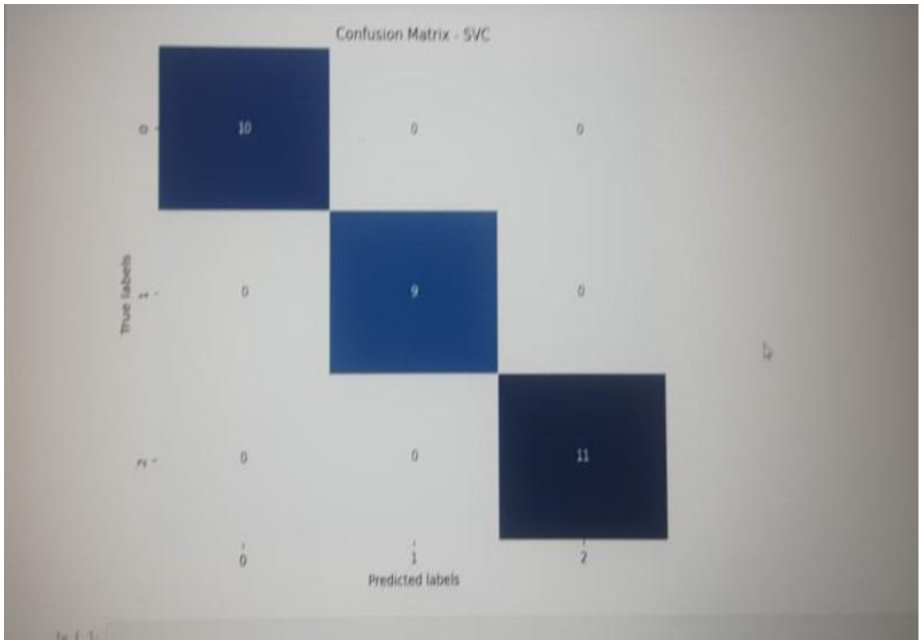
Multi-Layer-Perceptron Classification Report:

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Folk | 0.69 | 0.70 | 0.69 | 263 |
| Rock | 0.89 | 0.87 | 0.88 | 1193 |
| Classical | 0.87 | 0.92 | 0.89 | 72 |
| Electronic | 0.80 | 0.85 | 0.82 | 629 |
| Hip-Hop | 0.81 | 0.73 | 0.77 | 277 |
| accuracy | | | 0.83 | 2434 |
| macro avg | 0.81 | 0.81 | 0.81 | 2434 |
| weighted avg | 0.83 | 0.83 | 0.83 | 2434 |

Support Vector Machine Classification Report:

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Folk | 0.81 | 0.76 | 0.79 | 263 |
| Rock | 0.91 | 0.91 | 0.91 | 1193 |
| Classical | 0.97 | 0.89 | 0.93 | 72 |
| Electronic | 0.82 | 0.91 | 0.86 | 629 |
| Hip-Hop | 0.90 | 0.76 | 0.82 | 277 |
| accuracy | | | 0.87 | 2434 |
| macro avg | 0.88 | 0.85 | 0.86 | 2434 |
| weighted avg | 0.88 | 0.87 | 0.87 | 2434 |

Fig 6.2: Dataset


Confusion Matrix - SVC
True labels
Predicted labels
10
0
0
0
9
0
0
0
11
0
1
2
0
1
2

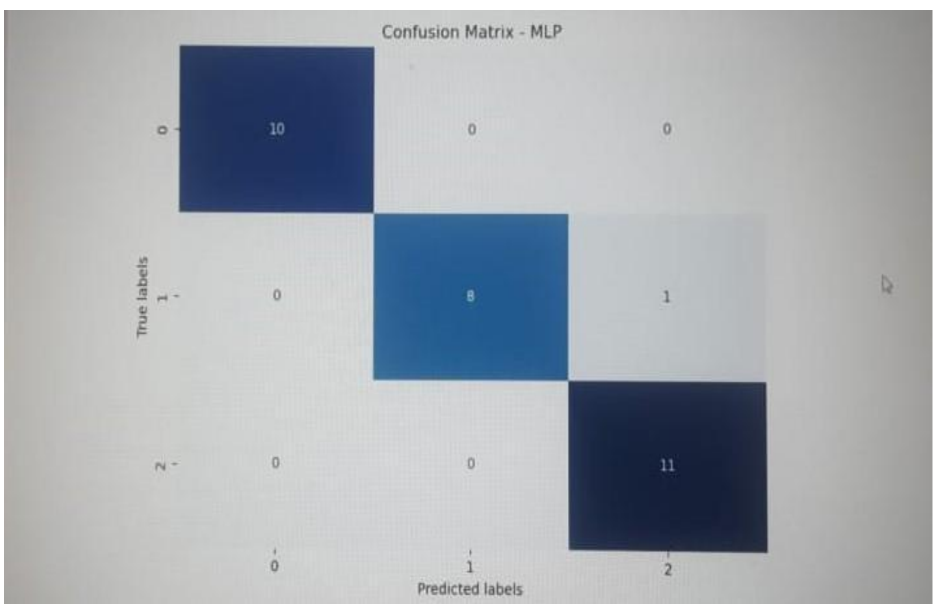

Confusion Matrix - MLP
True labels
Predicted labels
10
0
0
0
8
1
0
0
11
0
1
2
0
1
2

Fig 6.3

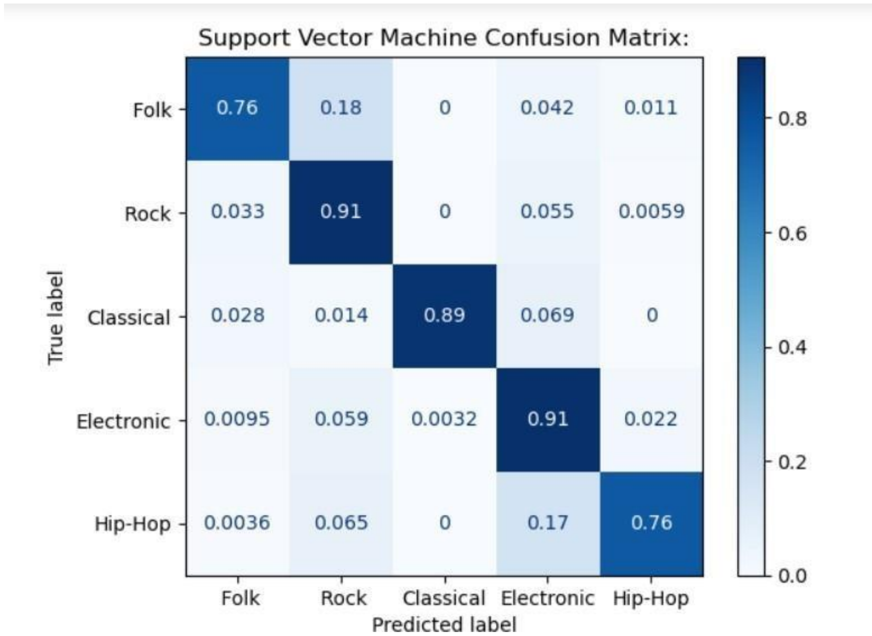

Support Vector Machine Confusion Matrix:
True label
Folk
Rock
Classical
Electronic
Hip-Hop
0.76 0.18 0 0.042 0.011
0.033 0.91 0 0.055 0.0059
0.028 0.014 0.89 0.069 0
0.0095 0.059 0.0032 0.91 0.022
0.0036 0.065 0 0.17 0.76
Folk Rock Classical Electronic Hip-Hop
Predicted label
0.0
0.2
0.4
0.6
0.8

Fig 6.4

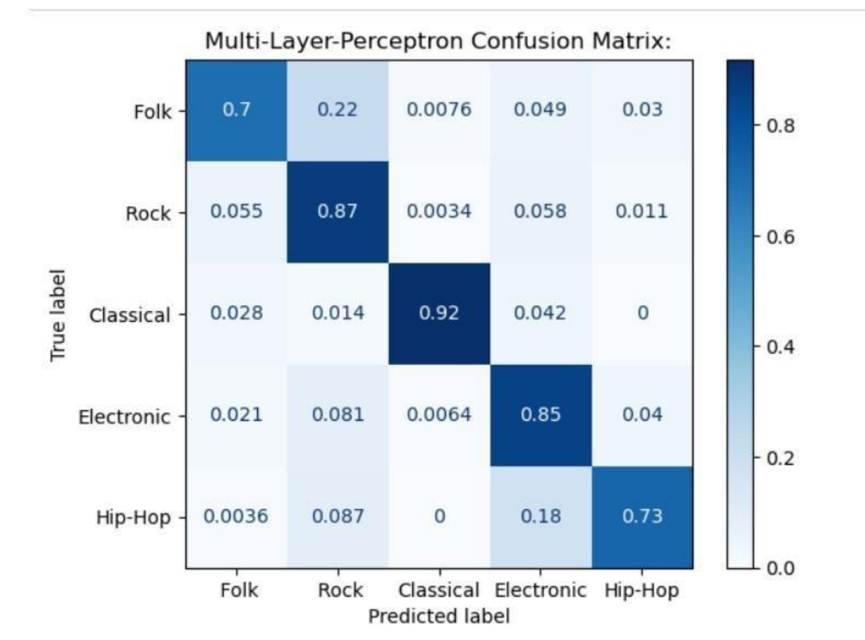

Multi-Layer-Perceptron Confusion Matrix:
True label
Folk
Rock
Classical
Electronic
Hip-Hop
0.7 0.22 0.0076 0.049 0.03
0.055 0.87 0.0034 0.058 0.011
0.028 0.014 0.92 0.042 0
0.021 0.081 0.0064 0.85 0.04
0.0036 0.087 0 0.18 0.73
Folk Rock Classical Electronic Hip-Hop
Predicted label
0.0
0.2
0.4
0.6
0.8

Fig 6.5

## CHAPTER 7

# CONCLUSION

In conclusion, the task of music genre classification is a multifaceted endeavor that combines elements of signal processing, machine learning, and domain expertise. Through the use of advanced algorithms and feature extraction techniques, researchers and practitioners have made significant strides in automating the classification process, thereby enabling a wide range of applications from recommendation systems to content organization. However, despite these advancements, challenges such as label ambiguity, cultural variability, and the inherent subjective nature of genre classification persist. Future research efforts should aim to address these challenges by leveraging novel approaches such as deep learning architectures, incorporating contextual information, and fostering interdisciplinary collaborations. By doing so, we can further refine and improve the accuracy and robustness of music genre classification systems, ultimately enhancing our understanding and appreciation of the diverse musical landscape. The combination of CNN, SVM and Random Forest has been proved to powerful technique formusic genre classification, as it leveraged the strengths of all three models. CNN extracted relevant features from the audio signals, while the SVM and Random Forest effectively separated different genresusing those features.

# REFERENCES


1. Campbell, Tzanetakis, George, and Perry Cook. "Musical genre classification of audio signals." IEEE Transactions onspeech and audio processing 10.5 (2002): 293-302.

2. Liang, Kehuang, and Liang Lin. "Music genre classification by learning from heterogeneous data." IEEETransactions on Multimedia 15.3 (2013): 610-621.

3. Tieleman, Tijmen, and Geoffrey Hinton. "Using very deep autoencoders for content-based audio retrieval."Advances in Neural Information Processing Systems. 2009.

4. Dieleman, Sander, et al. "Audio-based music classification with a pretrained convolutional network." 2011 12thInternational Society for Music Information Retrieval Conference (ISMIR). IEEE, 2011.

5. Pons, Jordi, et al. "Experimenting with musically motivated convolutional neural networks." 2016 IEEEInternational Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2016.

6. Sturm, Bob L. "A survey of evaluation in music genre recognition." Journal of New Music Research 33.3(2004): 239-266.

7. Cheng, Hao, et al. "A survey of music information retrieval systems." Journal of Intelligent InformationSystems 41.3 (2013): 355-381.

8. McFee, Brian, et al. "The librosa audio analysis library." Proceedings of the 14th python in scienceconference. 2015.

9. Humphrey, Eric J., et al. "Feature learning and deep architectures: new directions for music informatics." Journal of Intelligent Information Systems 41.3 (2013): 461-481.

10. Zhang, Yin, et al. "Automatic tagging using deep convolutional neural networks." 2016 IEEE International.