IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.301

Volume 3, Issue 13, May 2023

Music Genre Classification

Anish Umale, Mehul, Prajwal Bhandw, Sahil Bagwan, Prof. S. M. Patil

Department of Computer Engineering

Sinhgad College of Engineering, Pune, India

Abstract: *The music industry has undergone major changes from its conventional existence and in the form of music created in the last few years. The ever-growing customer base has also increased the market for different music styles. Music not only brings the individuals together, but also provides insight for various cultures. Therefore, it is essential to classify themusic according to the genres to satisfy the needs of the people categorically. The manual ranking of music is a repetitive, lengthy task and the duty lies with the listener.*

Keywords: *music.*

I. INTRODUCTION

With the growth of online music databases and easy access to music content, people find it increasingly hard to manage the songs that they listen to. One way 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. In this project we explore the application of machine learning (ML) algorithms to identify and classify the genre of a given 1 audio file.

Background and basics

One way of categorizing and organizing music 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.

A music genre is a conventional category that predicts the genre of music belonging to tradition or a set of conventions. Categorizing music files according to their genre is a challenging task in the area of music information retrieval. This Music Genres Classification System will detect the music from the audio file. Once the music is detected, the system will further continue to classification. As a result, the system will display the music genre. So, for this system, there are a predefined set of music genres that the system will classify.

Advanced music databases are continuously achieving a reputation in relations to specialized archives and private sound collections. Due to improvements in internet services and network bandwidth there is also an increase in the number of people involved with the audio libraries. But with a large music database the warehouses require an exhausting and time consuming work, particularly when categorizing audio genres manually.

Music has also been divided into Genres and subgenres not only on the basis of music but also on the lyrics as well [2]. This makes classification harder. To make things more complicated the definition of music genre may have very well changed over time [3]. For instance, rock songs that were made fifty years ago are different from the rock songs we have today.

II. LITERATURE SURVEY

Music Genre Classification using Machine Learning

Authors : Anirudh Ghildiyal; Komal Singh; Sachin Sharma, IEEE (2022),"

Abstract: The music industry has undergone major changes from its conventional existence and in the form of music created in the last few years. The ever-growing customer base has also increased the market for different music styles. Music not only brings the individuals together, but also provides insight for various cultures. Therefore, it is essential to classify the music according to the genres to satisfy the needs of the people categorically. The manual ranking of music is a repetitive, lengthy task and the duty lies with the listener. This research work has trained and compared the proposed models on GTZAN dataset, where most of the models were audio file trains, while a few of the models were trained on the spectrogram.[7]

Music Genre Classification
Authors : Jitesh Kumar Bhatia; Rishabh Dev Singh; Sanket Kumar, IEEE (2021)
Abstract: Music plays a vital role in people's lives. Music unites like-minded people and is the glue that binds communities together. Communities can be identified by the type of songs that are composed, or listened to. In this project, we will build an in-depth learning project to automatically distinguish genres of music from various audio files. We will classify these audio files using features at a low frequency level and time zone. In this project, we have built several segmentation models and trained them on the GTZAN database. Database has 1000 type audio tracks and the time duration of each track is 30 seconds. It contains 10 types of music genres; each music genre contains 100 tracks. All the tracks are in. wav format.[5]

Support vector machine active learning for music retrieval
Authors : Michael I. Mandel, Graham E. Poliner & Daniel P. W. Ellis, IEEE (2021)
Abstract: The paper discusses searching and organizing growing digital music collection and requires a computational model of music similarity. This paper describes a system for performing flexible music similarity queries using SVM active learning. We evaluated the success of our system by classifying according to mood and style (from an online music guide) and by the performing artist. In comparing a number of representations for songs, we found the statistics of mel-frequency cepstral coefficients to perform best in precision at comparisons. We also show that by choosing training examples intelligently, active learning requires half as many labeled examples to achieve the same accuracy as a standard scheme.[3]

Musical genre classification using support vector machines
Authors : Changsheng Xu, Namunu C. Maddage, Xi Shao, Fang Cao IEEE, (2020)
Abstract: Automatic musical genre classification is very useful for music indexing and retrieval. In this paper, an efficient and effective automatic musical genre classification approach is presented. A set of features is extracted and used to characterize music content. A multi-layer classifier based on support vector machines is applied to musical genre classification. Support vector machines are used to obtain the optimal class boundaries.This paper gives us the idea regarding the use of Support vector machines which are very useful for the implementation of classification.Support vector machine applications in classification as well as regression.[6]

Music Genre Classification using Support Vector Machine
Authors : Eamin Chaudary; Sumair Aziz; Muhammad Umar Khan, IEEE (2020)
Abstract: Classification of music classes is very tricky in the field of music information retrieval (MIR). In this article, a novel approach is proposed first, the audio signal is denoised and region of interest is extracted using time and frequency domain features are extracted then linear Support Vector Machine (SVM-L) is trained to get the best accuracy of 94.0% with 5 genres i.e. blues, classic, hip-hop, metal, and pop.[3]

A Hybrid Model For Music Genre Classification Using LSTM And SVM
Authors : Prasenjeet Fulzele, Rajat Singh, Naman Kaushik,Kavita Pandey, IEEE (2020) Abstract: With today's cutting edge technology and intractable access to voluminous data files via the internet, it is important to meet the computational needs of every user. Machine learning is one such growing branch of artificial intelligence that has made such demands of the users viable. Machine learning models are paving the way for classification techniques such as in music genre classification, and have shown to be efficient in predicting classes to a great extent. To exploit the time dependent nature of the dataset Long Short-Term Memory (LSTM) Neural Network is used for music genre classification and combined with Support Vector Machine (SVM) classifier to enhance its performance. The hybrid model of these two classifiers

resulted in an increase in the accuracy of prediction of the individual models. This hybrid model is imposed on the GTZAN music dataset and is compared with the results of standalone models of LSTM and SVM.[4]

2.1 Problem definition

In this project with the help of machine learning we are going to classify the music into one of the music genres Classical, Disco, Blues, Hiphop, Metal, Pop, Jazz, Country, Rock etc. The main task of Genre classification is to predict music genre using the audio signal. The system will classify the given set of music into any 10 genres and the dataset which is used to train the system is GTZAN dataset. There are different types of dataset but we are going to use GTZAN dataset as it is an open source dataset and it uses less computational power, that is it can be used easily to train the system.

2.2 Scope statement

Many companies nowadays use music classification, either to be able to place recommendations to their customers (such as Spotify, Soundcloud), or simply as a product (for example Shazam).The main purpose of the project Music Genre Classification is to predict music genre using the input audio signal.

- Research: As music evolves over the year this project can be used for research about various feature extractions.
- Recommendation System: Many music systems and artists nowadays rely on recommendation systems to recommend new and upcoming music to their users this project can help in the same.
- End User application: People usually prefer to base their playlists on genres and moods of the song and an end user application based on this project can be helpful in achieving the same.

We will design a machine learning model which uses SVM algorithm for music genre classification based on data given by the user as well as the input provided by the User.

2.3 Organization Of Project Report

The report is divided into 4 chapters, Chapter 1 gives a brief introduction to the project. Chapter 2 discusses project planning and management that covers system overview and functional and non-functional requirements. Chapter 3 discusses analysis and design. And to conclude Chapter 4.

III. PROJECT PLANNING AND MANAGEMENT

Technology advancement and digitization were not always regarded as partners in the evolution of the arts industry. This viewpoint and mindset are about to shift. With their involvement in various

Introduction

This chapter covers the project planning and management details. It also covers System Requirement specifications. SRS is considered as the base for the effort estimations and project scheduling.

System Requirement Specification (SRS)

System Overview

This project will be a standard machine learning model which will allow users to upload a
.mp3/.wav music file and let them know what genre like Blues, Classical, Country, Dance, Jazz, Latin, Pop, R&B, Rap, or Rock their music belongs to. This model can also be used as an end user application to classify music genres of people's music files and recommend similar kinds of music.

Functional Requirements

Focusing on the domain of software engineering, a functional requirement refers to the set of requirements for the system to function as expected. Functional requirements can refer to requirements pertaining to necessary calculations, necessary technical details, vital data manipulation and processing among other explicit functionalities that define what a system is intended to undertake.

In this project, the system aims to perform the following tasks:

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 13, May 2023

- Accept standard input datasets which is the GTZAN dataset.
- Convert the audio file into .mp3/.wav format for classification purpose.
- Extract the features of the audio file and store it in the form of .csv file.
- Use SVM for classification of input dataset into different music genres.

Non-Functional Requirements

Performance

The performance of the system should be high when executing the user's input and should be able to provide feedback or response within a short time span, usually 5 to 10 seconds for highly complicated tasks and 2 to 3 seconds for less complicated tasks.

Accessibility

The system should be platform independent so that it is easily accessible by every user and the ease of using the system increases.

Maintainability

The system needs to be easily maintainable and improved upon if the need to make changes arises. The system should follow an iterative process of development and should be modular to allow maintainability.

Scalability

The system provides an interface for users to upload their music and get the genres of music for the same; this can provide for other areas of research that deal with audio signals. Further the same system can be used for finding the frequencies and features of different audio songs.This can be used on a large scale for understanding the user music preferences according to various parameters which includes region, language, festivals, occasions and culture.This data collected can be useful for various artist over the globe to know more about the music taste and preference.

Hardware Requirements

- Processor – Intel Core i3 and above.
- RAM – 4GB DDR and above.
- Secondary Storage – Preferably 10GB and above.
- Input device – Standard keyboard and Mouse.
- Output device – VGA and High-Resolution Monitor

Software Requirements

- Librosa – Version 0.9.2
- Jupyter Notebook – Version 5.2.2
- Numpy – Version 1.11.3
- Pandas – Version 1.5.2
- Scikit-learn – Version 1.1.3
- Matplotlib – Version 3.6.2

Project Process Modeling

Agile project management is an iterative approach to delivering a project throughout its life cycle. Iterative or agile life cycles are composed of several iterations or incremental steps towards the completion of a project. Iterative approaches are frequently used in software development projects to promote velocity and adaptability since the benefit of iteration is that you can adjust as you go along rather than following a linear path. One of the aims of an agile or iterative approach is to release benefits throughout the process rather than only at the end. At the core, agile projects should exhibit central values and behaviors of trust, flexibility, empowerment and collaboration.

IV. ANALYSIS & DESIGN

This chapter covers the analysis and design of the considered system.

Idea Matrix

Idea	Feature	Modules Affected
Give users the ability to upload a music file.	Provide a web/program interface foruploading a music file.	Graphical User Interface
Accept .mp3 or .wav file.	Provide option to allow users to upload a .mp3 or a .wav file	System Architecture
Extract features from the audio file.	Extract various features like MFCC, Chroma frequency, Tempo, Spectral centroid from the audio file.	Testing/Output dataset
Implement an algorithm for genre classification	Use SVM for classification of inputdataset into different music genres.	ML model

Table 1 : Idea Matrix

V. ALGORITHM/METHODOLOGY

5.1 Feature Extraction

Mel-Frequency cepstral coefficients(MFCC)

Human perception of the frequency content of sounds does not follow a linear scale but uses a logarithmic distribution. Mel-frequency cepstral coefficients(MFCCs) are based onthe spectral information of a sound, but are modeled to capture the perceptually relevant parts of the auditory spectrum.[3]

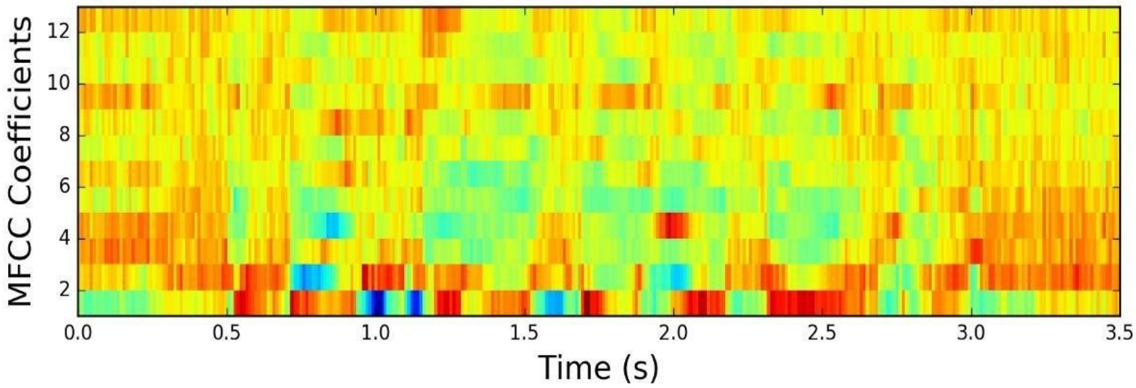


Fig.1 MFCC Coefficients for Rock music

Spectral centroid

The spectral centroid is a measure used in digital signal processing to characterize a spectrum. It indicated where the "center of mass" of the spectrum is located. Perceptually,it has a robust connection with the impression of "brightness" of a sound. It is calculated as the weighted mean of the frequencies present in the signal, determined using a Fourier transform with their magnitudes as the weights.The spectral centroid indicates at which frequency the energy of a spectrum is centered upon. This is like a weighted mean:

$$f_c = \frac{\sum_k S(k) f(k)}{\sum_k S(k)}$$

where S(k) is the spectral magnitude at frequency bin k, f(k) is the frequency at bin k

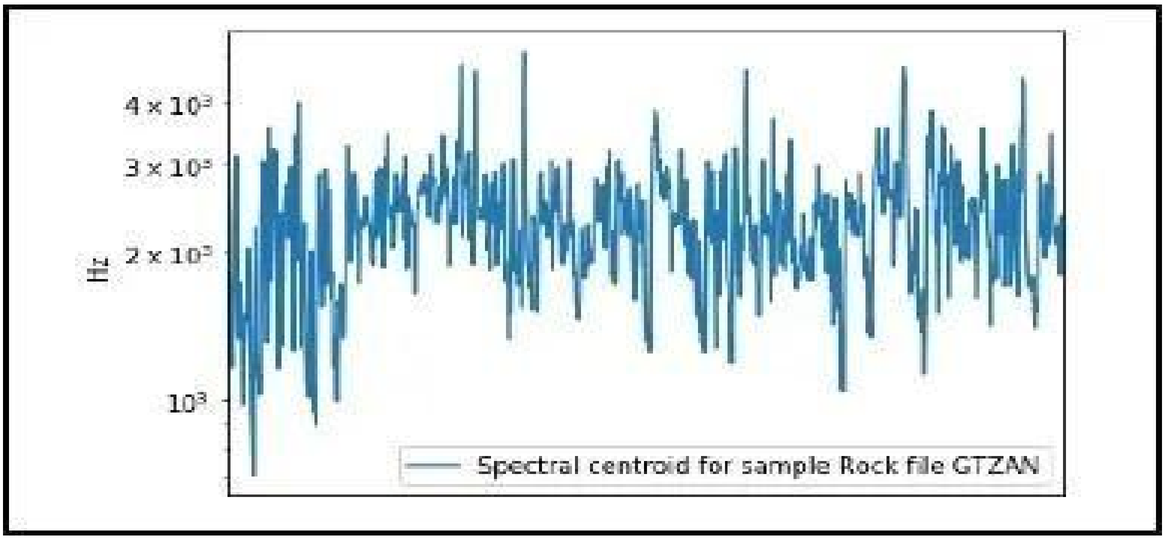


Fig.2 Spectral centroid for Rock music

Zero Crossing Rate

In the context of discrete-time signals, a zero crossing is said to occur if successive samples have different algebraic signs. The rate at which zero crossing occurs is a simple measure of the frequency content of a signal.[8] This average zero crossing rate gives a reasonable way to estimate the frequency of wave size. The number of zero crosses is alsoa useful feature in music analysis. Zero Crossing Rate is suitable for narrowband signals, but music signals include both narrowband and broadband components. Therefore, the short time zero crossing rate can be used to characterize music signals.[7]

Chroma Frequency

This is a vector that corresponds to the total energy of the signal in each of the 12 pitch classes. (C, C#, D, D#, E, F, F#, G, G#, A, A#, B). The Chroma vectors are then aggregated and their mean and standard deviation is taken.The underlying observation is that humans perceive two musical pitches as similar in color if they differ by an octave. Based on this observation, a pitch can be separated into two components, which are referred to as tone height and chroma. Assuming the equal-tempered scale, one considers twelve chroma values represented by the set {C, C♯, D, D♯, E , F, F♯, G, G♯, A, A♯, B} that consists of the twelve pitch spelling attributes as used in Western music notation.

Note that in the equal-tempered scale different pitch spellings such C♯ and D♭ refer to the same chroma. Enumerating the chroma values, one can identify the set of chroma values with the set of integers {1,2,...,12}, where 1 refers to chroma C, 2 to C♯, and so on.

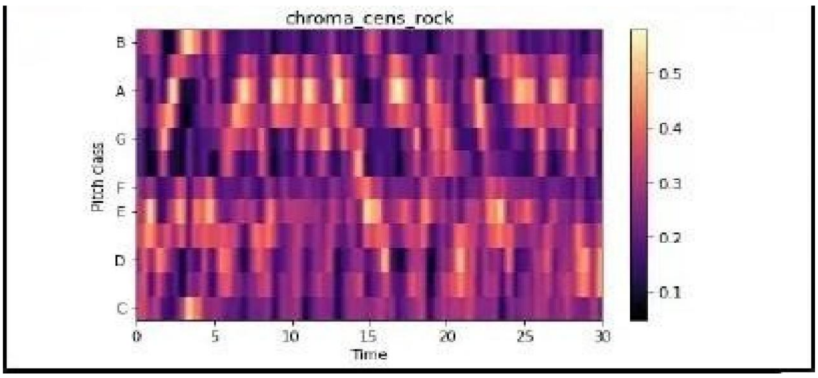


Fig.3 Chroma Frequency for Rock music

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 13, May 2023

Mel Spectrogram

Mel spectrogram is a spectrogram that is converted to a Mel scale. Then the concepts such as spectrogram and The Mel Scale are associated with it.[9]A spectrogram is a visualization of the frequency spectrum of a signal, where the frequency spectrum of a signal is the frequency range that is contained by the signal.Studies have shown that humans do not perceive frequencies on a linear scale.[6] We are better at detecting differences in lower frequencies than higher frequencies. For example, we can easily tell the difference between 500 and 1000 Hz, but we will hardly be able to tell a difference between 10,000 and 10,500 Hz, even though the distance between the two pairs are the same.The Mel scale mimics how the human ear works, with research showing humans don't perceive frequencies on a linear scale. Humans are better at detecting differences at lower frequencies than at higher frequencies. The name mel derives from melody and indicates that the scale is based on the comparison between pitches.The mel spectrogram remaps the values in hertz to the mel scale.The linear audio spectrogram is ideally suited for applications where all frequencies have equal importance, while mel spectrograms are better suited for applications that need to model human hearing perception. Mel spectrogram data is also suited for use in audio classification applications.

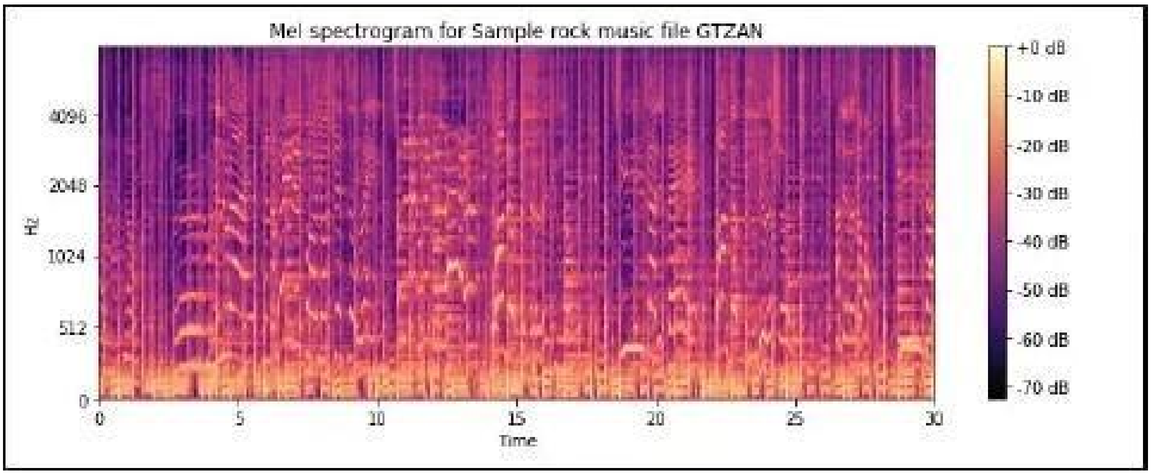


Fig.4 Mel Spectrogram for Rock Music

Supervised Learning

Supervised learning (SL) is a machine learning paradigm for problems where the available data consists of labelled examples, meaning that each data point contains features (covariates) and an associated label. The goal of supervised learning algorithms is learninga function that maps feature vectors (inputs) to labels (output), based on example input-output pairs. It infers a function from labeled training data consisting of a set of training examples.

Support Vector Machine(SVM)

Support vector machines are a set of supervised learning methods used for classification, regression, and outliers' detection. All of these are common tasks in machine learning.Youcan use them to detect cancerous cells based on millions of images or you can use them to predict future driving routes with a well-fitted regression model.There are specific types of SVMs you can use for particular machine learning problems, like support vector regression (SVR) which is an extension of support vector classification (SVC).The main thing to keep in mind here is that these are just math equations tuned to give you the most accurate answer possible as quickly as possible. SVMs are different from other classification algorithms because of the way they choose the decision boundary that maximizes the distance from the nearest data points of all the classes.

Working of SVM

Support vector machines also known as SVM is another algorithm widely used by machine learning people for both classification as well as regression problems but is widely used for classification tasks. It is preferred over other

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 13, May 2023

classification algorithms because it uses less computation and gives notable accuracy. It is good because it gives reliable results even if there is less data.

A simple linear SVM classifier works by making a straight line between two classes. That means all the data points on one side of the line will represent a category and the data points on the other side of the line will be put into a different category. This means there can be an infinite number of lines to choose from. What makes the linear SVM algorithm better than some of the other algorithms, like k-nearest neighbors, is that it chooses thebest line to classify the user's data points. It chooses the line that separates the data and is the furthest away from the closest data points as possible. A 2-D example helps to make sense of all the machine learning jargon. Basically, you have some data points on a grid. You are trying to separate these data points by the category they should fit in, but you don'twant to have any data in the wrong category.

That means you are trying to find the line between the two closest points that keeps the other data points separated. So, the two closest data points give you the support vectorsyou will use to find that line. That line is called the decision boundary. The decision boundary does not have to be a line. It is also referred to as a hyperplane because you can find the decision boundary with any number of features, not just two.There can be many hyperplanes that you can see but the best hyper plane that divides the two classes would bethe hyperplane having a large distance from the hyperplane from both the classes. That is the main motive of SVM to find such best hyperplanes.There can be different dimensions which solely depend upon the features we have

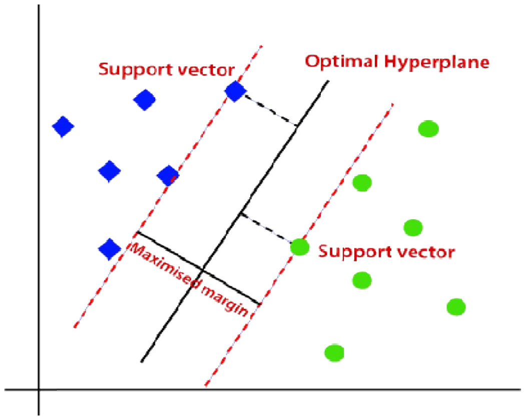


Fig.5 Working of SVM

Types of SVM

There are two different types of SVMs, each used for different things:

- **Simple SVM:** Typically used for linear regression and classification problems.
- **Kernel SVM:** Has more flexibility for non-linear data because you can add more features to fit a hyperplane instead of a two dimensional space.

Use of SVM

Music Genre Classification will be a standard machine learning model that can be used by any music lover. We may use the support vector learning algorithm to construct the following three types of learning machines (among others):

Polynomial learning machines

Radial-basis function networks

Two-layer perceptron's (i.e., with a single hidden layers)

That is, for each of these feed-forward networks we may use the support vector learning algorithm to implement the learning process using a given set of training data, automatically determining the required number of hidden units. Stated in another way: Whereas the back-propagation algorithm is devised specifically to train a multilayer perceptron, the support vector.

UML Diagrams

Class Diagram

The class diagram depicts a static view of an application. It represents the types of objects residing in the system and the relationships between them. A class consists of its objects, and also it may inherit from other classes. A class diagram is used to visualize, describe, document various different aspects of the system, and also construct executable software code.

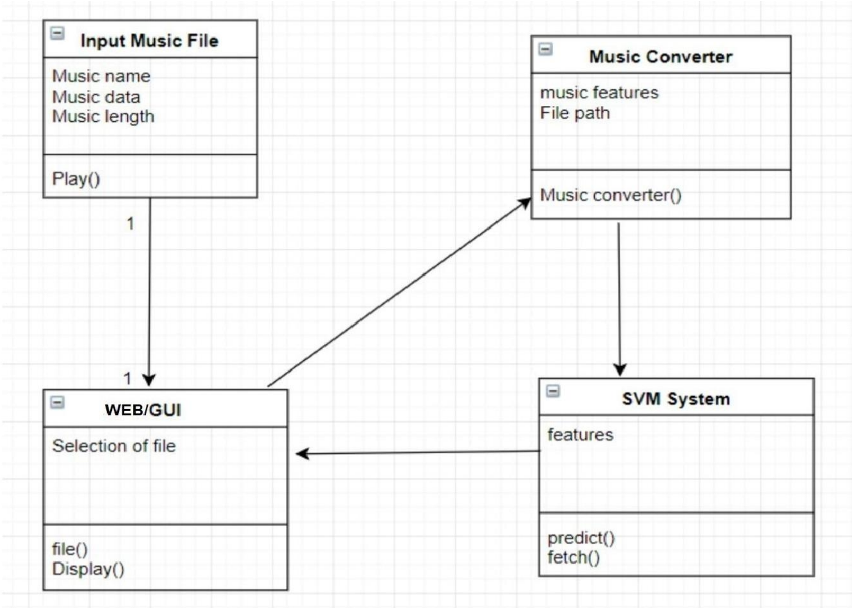


Fig.6 Class Diagram

Use Case Diagram

A use case diagram is used to represent the dynamic behavior of a system. It encapsulates the system's functionality by incorporating use cases, actors, and their relationships. It models the tasks, services, and functions required by a system/subsystem of an application. It depicts the high-level functionality of a system and also tells how the user handles a system.

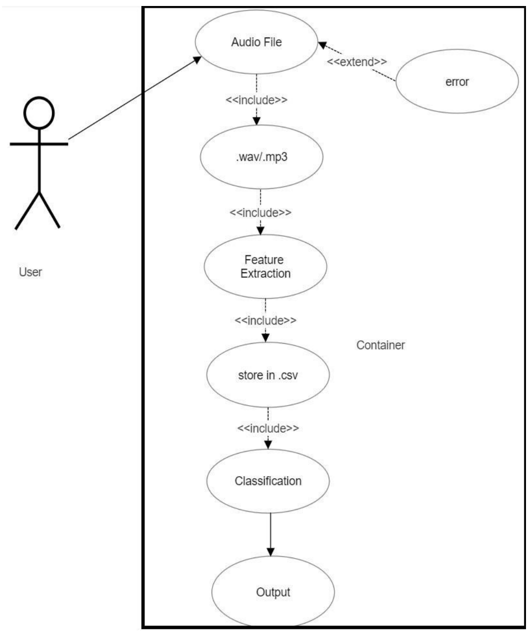
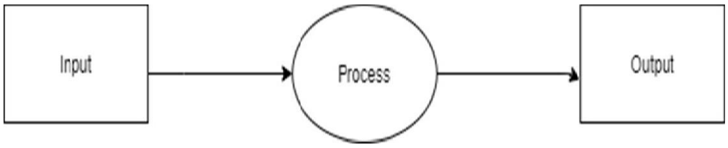


Fig.7 Use case Diagram

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 13, May 2023

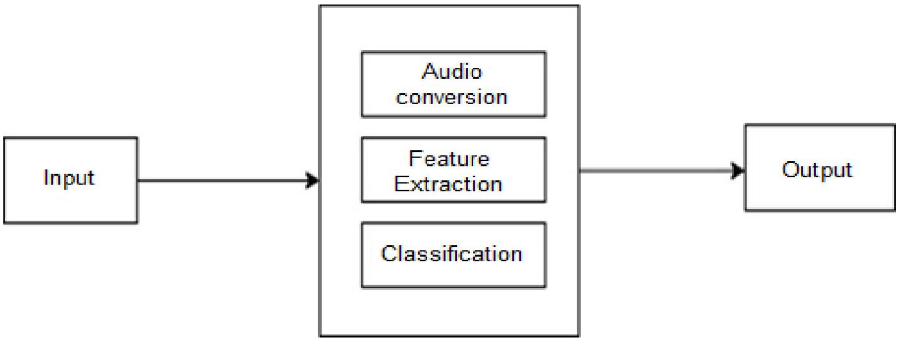
Data Flow Diagram

A data-flow diagram is a way of representing a flow of data through a process or a system (usually an information system). The DFD also provides information about the outputs andinputs of each entity and the process itself. A data-flow diagram has no control flow — there are no decision rules and no loops. Specific operations based on the data can be represented by a flowchart.



Data Flow Diagram (0)

Data Flow Diagram (1)



For each data flow, at least one of the endpoints (source and / or destination) must exist in a process. The refined representation of a process can be done in another data-flow diagram, whichsubdivides this process into sub-processes. The data-flow diagram is a tool that is part of structured analysis and data modeling. When using UML, the activity diagram typically takesover the role of the data-flow diagram. A special form of data-flow plan is a site-oriented data-flow plan.

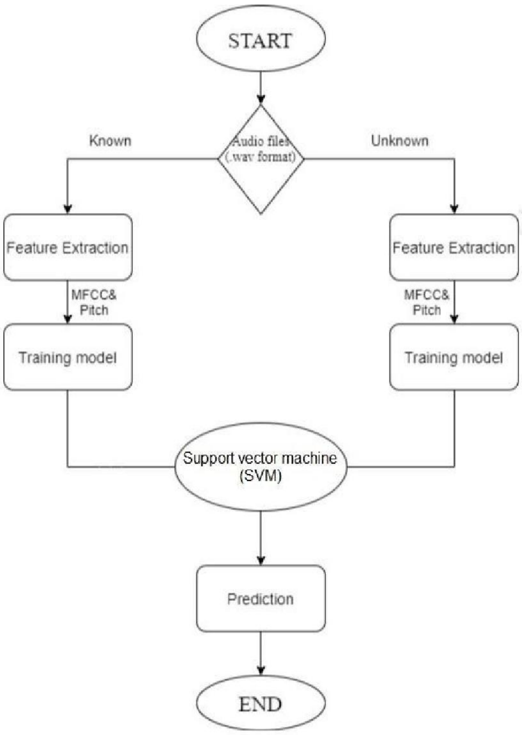


Fig.8 Data Flow Diagram

DOI: 10.48175/568

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 13, May 2023

Activity Diagram

Activity diagrams are graphical representations of workflows of stepwise activities and actionswith support for choice, iteration and concurrency.

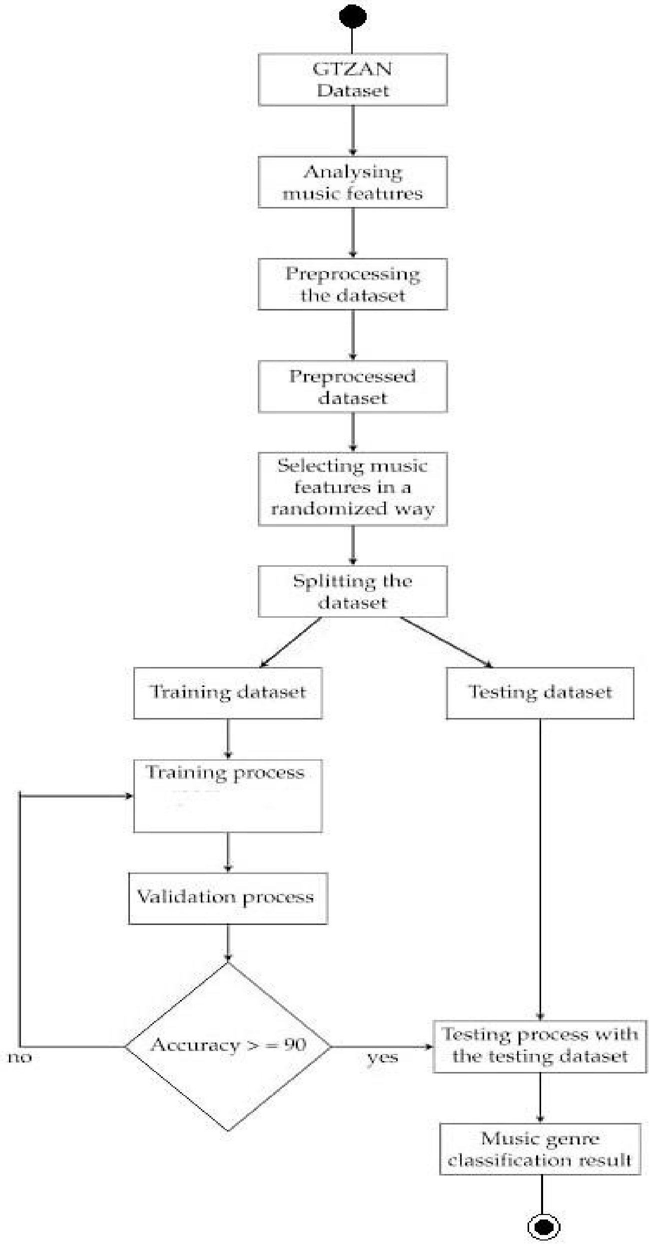


Fig.9 Activity Diagram

V. CONCLUSION

This project entailed securing a standardized dataset, deciding on an efficient model, training the model against the dataset, and making predictions using the same. with all these steps completed. We have successfully managed to implement a supervised learning algorithm to classify a given Test music file into its designated genre. The model produces an accuracy metric of 86% and is a decent result for a SVM model with a dataset of this scale. In conclusion the project has successfully achieved the objectives.

REFERENCES

[1] C. Joder, S. Essid, G. Richard, and S. Member, “Temporal Integration for Audio Classification With Application to Musical Instrument Classification,” IEEE Transactions on Speech and Audio Processing, 2022.

[2] Mi Tian, Gy ¨ orgy Fazekas, Dawn A. A. Black, Mark Sandler, On the use of the tempogram to describe audio content and its Application to music structural segmentation, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019.

[3] George Tzanetakis and Perry Cook. 2020. Musical genre classification of audio signals. IEEE Transactions on speech and audio processing 10(5):293-302

[4] Prasenjeet Fulzele, Rajat Singh, Naman Kaushik,Kavita Pandey, A Hybrid Model For Music Genre Classification Using LSTM And SVM IEEE 2020

[5] Changsheng Xu, Namunu C. Maddage, Xi Shao, Fang Cao.Musical genre classification using support vector machines. IEEE, 2020 Feature extractions https://en.wikipedia.org/wiki/Feature_extraction

[6] Anirudh Ghildiyal; Komal Singh; Sachin Sharma. Music Genre Classification using Machine Learning IEEE 2022

[7] Dan Ellis. 2007. Chroma features analysis and synthesis. Resources of Laboratory for the Recognition and Organization of Speech and Audio-Librosa.

[8] Steve Tjoa. 2017.Music information retrieval. https://musicinformationretrieval.com/spectral_features.html