Vision and Mission of the Institute

Vision

"To be a centre of excellence recognized nationally and internationally, in distinctive areas of engineering education and research, based on a culture of innovation and invention."

Mission

"BIET contributes to the growth and development of its students by imparting a broad based engineering education and empowering them to be successful in their chosen field by inculcating in them positive approach, leadership qualities and ethical values."

Vision and Mission of the Department

Vision

"To be a centre-of-excellence by imbibing state-of-the-art technology in the field of Computer Science and Engineering, thereby enabling students to excel professionally and be ethical."

Mission

1. Adapting best teaching and learning techniques that cultivates Questioning and Reasoning culture among the students.
2. Creating collaborative learning environment that ignites the critical thinking in students and leading to the innovation.
3. Establishing Industry Institute relationship to bridge the skill gap and make them industry ready and relevant.
4. Mentoring students to be socially responsible by inculcating ethical and moral values.

Program Educational Outcomes (PEOs)

PEO1	To apply skills acquired in the discipline of Computer Science and Engineering for solving societal and industrial problems with apt technology intervention.
PEO2	To continue their career in industry/academia or to pursue higher studies and research.
PEO3	To become successful entrepreneurs, innovators to design and develop software products and services that meets the societal, technical and business challenges.
PEO4	To work in the diversified environment by acquiring leadership qualities with effective communication skills accompanied by professional and ethical values.

Program Specific Outcomes (PSOs)

PSO1	Analyze and develop solutions for problems that are complex in nature by applying the knowledge acquired from the core subjects of this program.
PSO2	Ability to develop Secure, Scalable, Resilient and distributed applications for industry and societal requirements.
PSO3	Ability to learn and apply the concepts and construct of emerging technologies like Artificial Intelligence, Machine learning, Deep learning, Big Data Analytics, IoT, Cloud Computing, etc for any real time problems.

ABSTRACT

The acoustic features of music have been extracted by using digital signal processing techniques and then using neural network, music genre classification have been done. We propose to use a model which means to make use of the GTZAN database for data analysis and modelling. The dataset uses images of spectrograms generated from songs as the input into a neural net model to classify the songs into their respective musical genres. The objective of our project work is to implement supervised learning technique Artificial Neural Networks for classifying musical categories. Thus, comparing music classifiers accuracy for datasets of different nature. Also, the classification accuracy of genre classifiers on the varying number of modifiers and layers is analysed.

CONTENTS

CHAPTER 1

INTRODUCTION

Music plays a very important and impacting role in people's lives. Music brings like-minded people together and is the glue that binds groups and communities together. The widespread usage of the Internet has brought about significant changes in the music industry as well as leading to all kinds of change. Examples of these developments being the widespread usage of online music listening and sales platforms, control of music copyright, classification of music genre, and music recommendations. Today, with the advancement of music broadcast platforms, people can listen to music at any time and at any time and can reach millions of songs through various music listening platforms such as Spotify, Sound Cloud, iTunes, Saavn etc.

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

1.1 Audio processing

Audio Processing unit consists of analysing the time domain features such as Tempo, Amplitude, etc of the data along with frequency domain features of the data. This analysis helps us to identify the defining features of an audio wave and understand the required methodology for features which influence the distinction of any musical genre.

1.1.1 AUDIO ANALYSIS (TIME & FREQUENCY DOMAIN)

Sound is typically represented in the form of an audio signal having parameters such as frequency, bandwidth, decibel, etc. A typical audio signal can be expressed in the time domain as a function of Amplitude and Time.

Time Domain

A time domain analysis is the analysis of physical signals, mathematical functions, or time series of any data, with reference to time. Also, in the time domain, the signal or function's value is found for all real numbers at numerous separate instances in the case of discrete time or in the case of continuous time. A time domain graph can show how a signal changes with respect to time, whereas a frequency domain graph will show what proportion of the signal lies within each given waveband over a range of frequencies.

Frequency Domain

Frequency domain is the analysis of signals or mathematical functions, with reference to frequency. As mentioned earlier, a time domain graph shows the changes in a signal over a period of time, and frequency domain shows what proportion of the signal exists within a given waveband over a range of frequencies. Also, a frequency domain representation can include information on the phase shift that must be applied to each sinusoid to be able to recombine the frequency components to recover the original time signal.

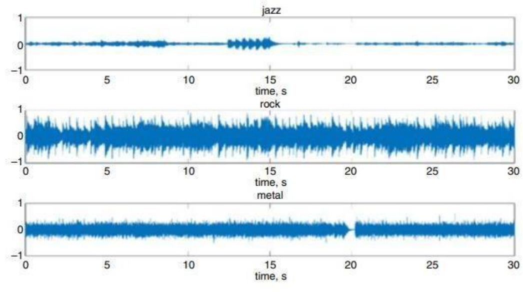

Fig 1.1.1 Example time domain representation of Jazz, Rock and Metal

1.2 Feature Extraction

Feature Extraction unit consists of gathering all the required and characteristic features of audio and consolidating them all in an ordered format for further ad hoc data analysis and identifying any anomalies or discrepancies in the data. This ordered format (.csv format) has been used for predictive modelling aspect of the Neural Network.

1.2.1 Mel Frequency Cepstral Coefficients

These are state-of-the-art features used in automatic speech and speech recognition studies. There are a set of steps for generation of these features:

- Since the audio signals are constantly changing, first we divide these signals into smaller frames. Each frame is around 20-40 ms long
- Then we try to identify different frequencies present in each frame
- Now, separate linguistic frequencies from the noise
- To discard the noise, it then takes discrete cosine transform (DCT) of these frequencies. Using DCT we keep only a specific sequence of frequencies that have a high probability of information.

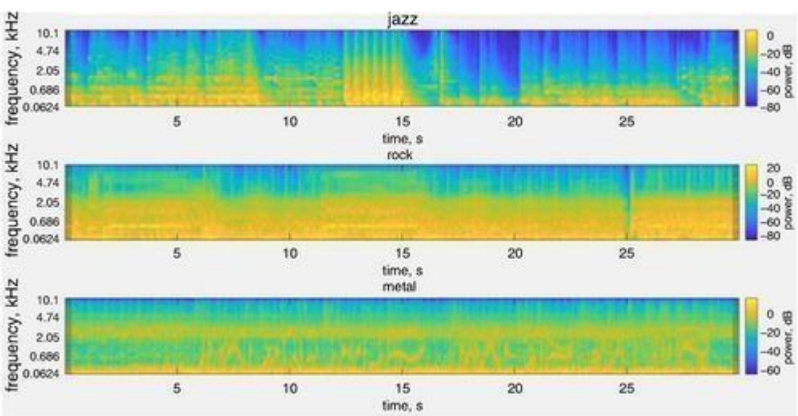

Fig 1.2.1 Feature Extraction of an audio

1.3 Types of music genre

A music genre is a classification system that classifies music into different styles. It's the art of incorporating instrumental and vocal tones in a structured manner that gives the music its distinctive character. The word genre is used in other forms of art, including literature, television, cinema, and other artistic creation types. It combines pieces of work that fit under a specific category after analyzing and highlighting the most distinctive elements.

Ex;

- Blues
- Classical
- Country
- Disco
- Hip hop
- Jazz
- Metal
- Pop
- Reggae
- Rock

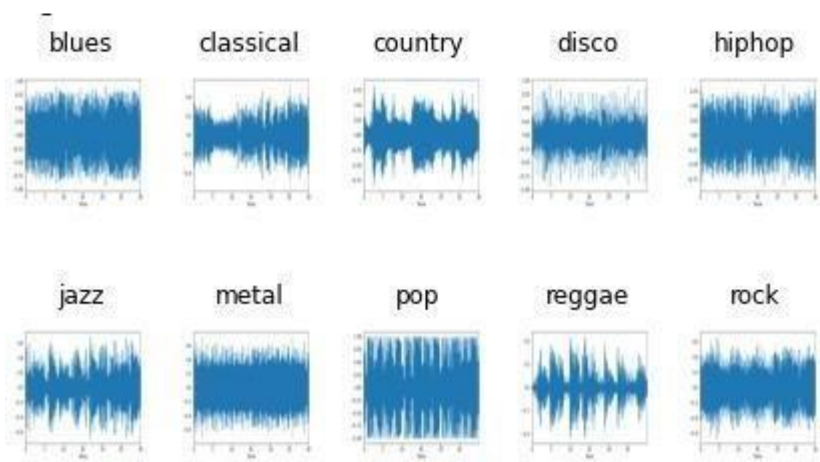

Fig 1.3.1 Different music genre's

CHAPTER 2

LITERATURE SURVEY

Music genre labels are useful to organize songs, albums, and artists into broader groups that share similar musical characteristics. 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 of the ways to categorize and organize songs is based on the genre, which is identified by some defining characteristics of the music. Music genre classification has been a widely studied area of research since the early days of the Internet. Musical genres have no strict definitions and boundaries as they arise through a complex interaction between the public, marketing, historical, and cultural factors. This observation has led some researchers to suggest the definition of a new genre classification scheme purely for the purposes of music information retrieval [1][2]

According to Aucouturier and Pachet, 2018 [3] genre of music is possibly the best general information for the music content clarification. 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. Tzanetakis and Cook (2019)[4] addressed this problem with supervised machine learning approaches such K-Nearest neighbour classifiers. More recent deep learning approaches take advantage of visual representations of the audio signal in form of spectrograms. These visual representations are used as input to Convolutional Neural Networks (CNNs)[5]

In Lidy and Rauber (2020),[6] the authors discuss the contribution of psycho-acoustic features for recognizing music genres, especially the importance of STFT taken on the Bark Scale (Zwicker and Fastl, 2019). Mel-frequency cepstral coefficients (MFCCs), spectral contrast and spectral roll-off were some of the features used by (Tzanetakis and Cook, 2020)[1]. A combination of visual and acoustic features are used to train SVM and AdaBoost classifiers in Nanni et al. (2016).

With the recent success of deep neural networks, a number of studies apply these techniques to speech and other forms of audio data (AbdelHamid et al., 2017; Gemmeke et al., 2019[8]. Representing audio in the time domain for input to neural networks is not very straight-forward because of the high sampling rate of audio signals. However, it has been addressed in Van Den Oord et al. (2017) [9] for audio generation tasks. A common alternative representation is the spectrogram of a signal which captures both time and frequency information. Spectrograms can be considered as images and used to train convolutional neural networks (CNNs) (Wyse, 2019) [10]. A CNN was developed to predict the music genre using the raw MFCC matrix as input. In Lidy and Schindler (2019) [11], a constant Q-transform (CQT) spectrogram was provided as input to the CNN to achieve the same task.

Current models have only been focused on CNN models which involve using images (Spectrogram graphs) as the input data for learning attributes of genres. CNN models require higher computational power to process images instead of numerical data. Research over models utilising only key peak features has not been done before extensively. We will be using only the key characteristic features of the audio file instead of the entire spectrogram for data processing and learning

2.1 Literature Review Summary Table

The acoustic features of music have been extracted by using digital signal processing techniques and then using neural net, music genre classification have been done. We use the GTZAN database for data analysis and modelling. The dataset uses images of spectrograms generated from songs as the input into a neural net model to classify the songs into their respective musical genres.

Table 2.1. **Literature Review Summary Table**

Sl.no	Author	Year	Title	Summary
1	Elbir, A., and Aydin, N	2021	"Music Genre Classifier"	This System is a music genre classifier based on Acoustic features of music. This system has been designed to classify the genre of music basing upon the music features as well as the Manual listings of genre.
2	Shin, S.-H., Yun, H.-W., Jang,	2019	"Genre Classification Based On Acoustic Features"	This system has been designed to classify the genre of music basing upon the music features and also to classify the genre based on types of audio input.
3	Abdel Hamid & Gemmeke	2018	"Audio processing and classification"	Music genre labels are useful to organize songs, albums, and artists into broader groups that share similar musical characteristics. 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 of the ways to categorize and organize songs is based on the genre, which is identified by some defining characteristics of the music
4	Aucouturier and Pachet,	2018	"Music feature extraction and classification"	Musical genres have no strict definitions and boundaries as they arise through a complex interaction between the public, marketing, historical, and cultural factors
5	Lidy & Rauber	2017	"Classification based on visual and acoustic features"	A combination of visual and acoustic features are used to train SVM and AdaBoost classifiers

2.2 Existing System

Existing System is a music genre classifier based on Acoustic features of music(such as ,tempo, Instrumental , loudness). This system has been designed to classify the genre of music basing upon the music features as well as the Manual listings of genre.

2.3 Problem Statement

There are many genres of music and they are different from each other, resulting in people to have different preferences of music. so it is an important and up to date issue to classify music in music listening applications and platforms

2.4 Proposed System

The proposed System is a music genre classifier system based on signal processing Using Mel spectrogram and ANN classification. This system has been designed to perform better than the previous classifier Which used acoustic features to classify. The proposed System helps a Classification engine to Classify the music based on characteristics and features of music. The feature extraction will be done by Frequency and Time Domain Audio processing of the audio data.

2.5 Objectives

The main objectives of this project are:

- To develop a web-based application that can be used effectively to list and classify music genre accordingly.
- To evaluate and analyse the performance of the model for different classifier based of genres.
- To extract set of rules predicting the results into clusters of genres.s

CHAPTER 9

REQUIREMENTS

3.1 Software Requirements

The software required for the development of this project is:

- Environment : Google Colaboratory
- Programming Language : Python 3.
- Libraries : Tensorflow , Matplotlib.
- Front end : HTML, CSS, Java Script.
- Back end : Django Framework.
- Database : Postgresql.

3.2 Hardware Requirements

- Processor type : intel i5 or higher
- Hard Disk : 512 MB (min)
- RAM Size : 8GB (min)

CHAPTER 4

METHODOLOGY

Initially, each music in the data set is divided into six parts with a duration of 5 s. Mel spectrogram is generated from sampled each 5 s music and saved as an image. Then, this image is applied to the proposed Model for training. The Model which is shown as the last block in Fig. 3, is a type of CNN that new layers and artificial dropout features have been added to minimise validation error. Specifically, in our experiments, This System is designed to have three layers. Each layer consists of a two-dimensional convolution, an activation function (rectified linear unit), a two-dimensional maximum pooling operation and a dropout operation. After the training, the model is used for genre classification. Additionally, the last layer of the model named as Dense_2 is used as a feature vector of the test music samples for music genre classification, music similarity and music recommendation. We implemented classification algorithms

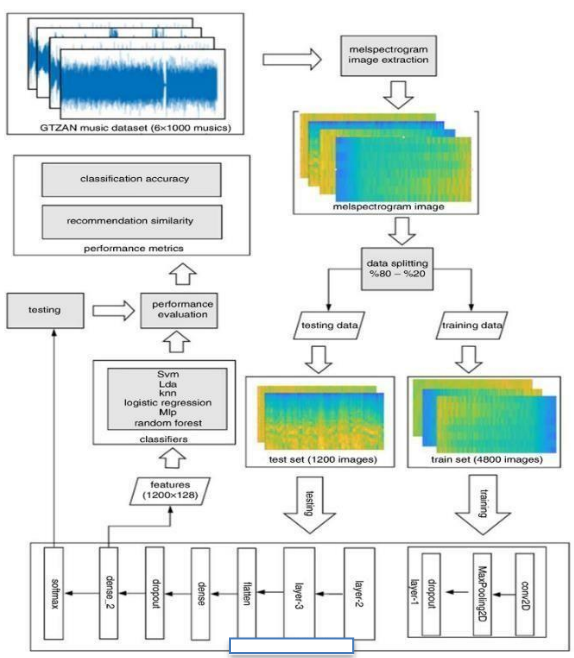

Fig 4.1: Training the Model

DATASET

We analyzed the features extracted from the GTZAN dataset and built different types of ensemble models to see how better we can differentiate one genre from another. Our Datasets contains 10 genres:

- Blues
- Classical
- Country
- Disco
- Hip hop
- Jazz
- Metal
- Pop
- Reggae
- Rock

We have 60 columns in our original dataset and we work with 10 of these for making our model.

- Name
- length of the file
- chroma shift mean
- chroma shift variance
- rms mean
- rms variance
- spectral centroid mean
- spectral centroid variance
- spectral bandwidth mean
- spectral bandwidth variance