# Music genre recognition using Deep Learning

A Project Report in partial fulfillment of the degree

**Bachelor of Technology**

in

**Computer Science &Engineering**

**By**


| | |
|---|---|
| 19UK1A0519 | A.Akhila |
| 19UK1A0531 | R.Pravalika |
| 19UK1A0523 | U.Sushmitha |
| 19UK1A0564 | B.Prasanna |


**Submitted to**


**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**
**VAAGDEVI ENGINEERING COLLEGE, BOLLIKUNTA, WARANGAL**
**(Affiliated to JNTUH, Accredited by NBA)**

## DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

## CERTIFICATE

This is to certify that the Project Report entitled "Music genre recognition" is a record of bonafide work carried out by the student(s) Akhila,Pravalika,Sushmitha,Prasanna bearing Roll No(s) 19UK1A019,19UK1A0531,19UK1A0523,19UK1A0564 during the academic year 2022-2022 in partial fulfillment of the award of the degree of Bachelor of Technology in **Computer Science Engineering by the Jawaharlal nehru technological univercity,hyderabad**

.

**Supervisor** **Head of the Department**

**External Examiner**

## ABSTRACT

In this paper,we have explored the use of artificial neural networks (ANNs) for -automatically detecting the genure of music.The challenge faced when hand classifying music is that it is highly dependent on the accuracy and experience of the person classifying it.the main objective of this paper is to build an arrangement which will reduce the burden and increase the accuracy of classifying the genre of music.we have trained our model on a novel dataset that reflects current trends in music and address the problem faced with existing datasets.

This model generates music genre from an input with an accuracy of around 90%

# Table of Contents

## 1. INTRODUCTION

The aim of this project was to compare machine learning algorithms in their ability to automatically classify song excerpts into the correct musical genre. A more fine-grained classification into musical subgenre usually requires some level of expertise on the part of the listener, and the greater the level of expertise, the more accurate the classification that can be made. The question is raised as to whether software can be written to perform genre classifications as well as humans, and what type of approaches work best. Users may want to discover tracks that are similar to those that they are already a fan of, and it makes sense for providers to have such metadata for their songs available, in order to facilitate this kind of discovery. Given the vast and ever-growing quantity of music available on the internet, and the need for services to manage increasingly large song databases, manual analysis and annotation of individual songs does not seem like a sustainable long-term solution. From an engineering perspective, despite the cost of developing and maintaining the system, long-term expenditures could be significantly reduced with the use of accurate, automated genre-recognition systems. Large streaming services are surely already aware of the prospects of automated genre recognition techniques, but questions remain as to the best way a music genre recognition system can be architected. This study compares several machine learning algorithms in their ability to classify songs accurately. First, a survey of existing literature is carried out, looking at how machine learning has been applied to problems of music classification and audio classification in general. Tools for extracting features from audio files are discussed, as well as techniques for classifying music based on other modalities such as song lyrics and album artwork. This is followed by a discussion of machine learning techniques fit for present purposes.

## 2. LITERATURE REVIEW

Music genre classification has been a widely studied area of research since the early days of the Internet. Tzanetakis and Cook (2002) addressed this problem with supervised machine learning approaches such as Gaussian Mixture model and k-nearest neighbour classifiers. They introduced 3sets of features for this task categorized as timbral structure, rhythmic content and pitch content. Hidden Markov Models (HMMs), which have been extensively used for speech recognition tasks, have also been explored for music genre classification (Scaringella and Zoia,2005;Soltauet al.,1998). Support vector machines (SVMs) with different distance metrics are studied and compared in Mandel and Ellis (2005) for classifying genre. In Lidy and Rauber (2005), the authors dis-cuss the contribution of psycho-acoustic features for recognizing music genre, especially the importance of STFT taken on the Bark Scale (Zwickerand Fastl,1999). Mel-frequency cepstral coefficients (MFCCs), spectral contrast and spectral roll-off were some of the features used by Tzane-takis and Cook,2002). A combination of visualand acoustic features are used to train SVM and AdaBoost classifiers in Nanni et al. (2016).With the recent success of deep neural net-works, a number of studies apply these techniquesto speech and other forms of audio data (Abdel-Hamid et al.,2014;Gemmeke et al.,2017). Rep-resenting audio in the time domain for input to neural networks is not very straight-forward be-cause of the high sampling rate of audio signals. However, it has been addressed in Van Den Oordet al. (2016) 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.

## 3. DESIGN:

### 3.1 Requirement Specifications (S/W & H/W)

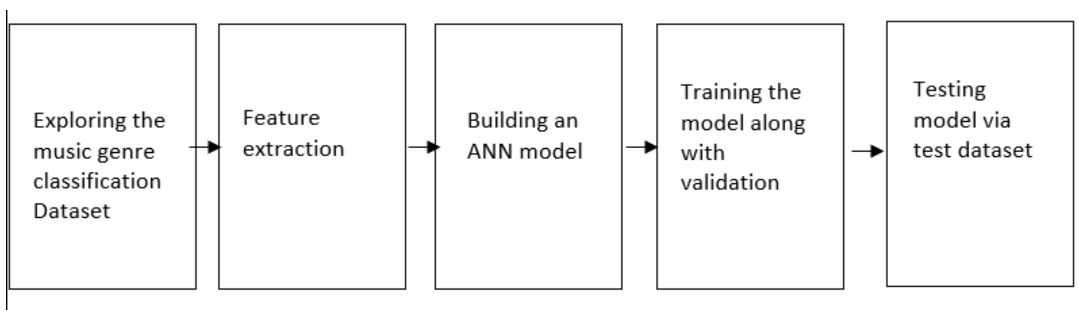
**Hardware Requirements**

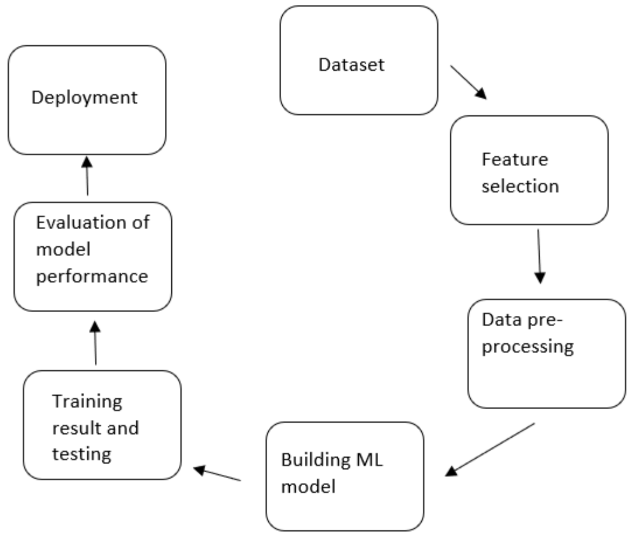
- ✓ **System** : Processor Intel(R) Core (TM) i5-8265U CPU @ 1.60GHz, 1800 MHz, 4 Cores, 8 Logical Processors
- ✓ **RAM** : 8 GB
- ✓ **Hard Disk** : 557 GB
- ✓ **Input** : Keyboard and Mouse
- ✓ **Output** : PC

**Software Requirements**

- ✓ **OS** : Windows 10/11
- ✓ **Platform** : Google Collaboratory
- ✓ **Deployment software** : Flask
- ✓ **Program Language** : Python

### 3.2 Flow chart

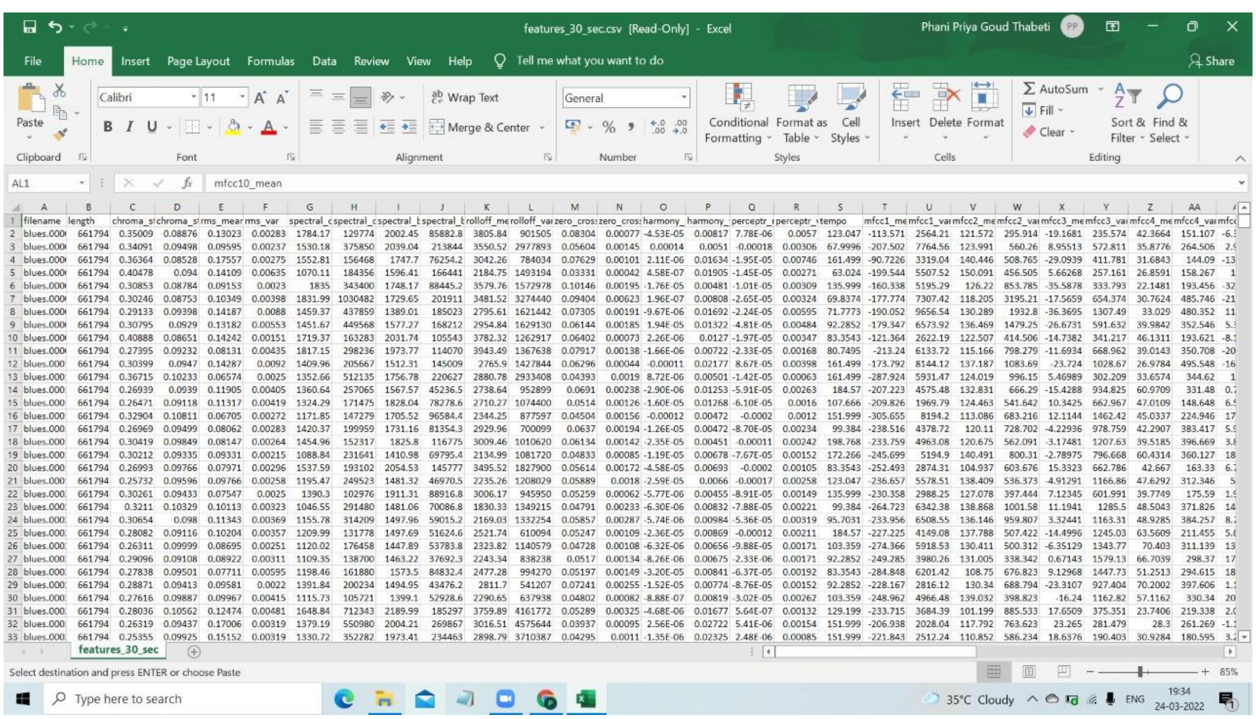





In the above flow chart we described our work flow in this project for developing deep learning model.

## 4. DATASET:

We collected data 1000 samples through a survey using Google forms Upon consideration, we finalized 1 input features to classify the music genre.



genres original - A collection of 10 genres with 100 audio files each, all having a length of 30 seconds (the famous GTZAN dataset, the MNIST of sounds)
images original - A visual representation for each audio file. One way to classify data is through neural networks. Because NNs (like CNN, what we will be using today) usually take in some sort of image representation, the audio files were converted to Mel Spectrograms to make this possible.
2 CSV files - Containing features of the audio files. One file has for each song (30 seconds long) a mean and variance computed over multiple features that can be extracted from an audio file. The other file has the same structure, but the songs were split before into 3 seconds audio files (this way increasing 10 times the amount of data we fuel into our classification models). With data, more is always better.

## 5. DATA PREPROCESSING:

### Feature Extraction

In order to represent the tracks numerically, 40 audio features were extracted from each track. These values represent the average over the entire track.
The 40 were extracted using LibROSA audio analysis library.
In total, this consisted chroma_stft, rms mean and variance, spectral centroid mean and variance, spectral bandwidth mean and variance, roll off means and variance, zero crossing rate mean and variance, harmonic mean and variance, per sector mean and variance, tempo.
'Tempo' was the only rhythmic feature extracted, and 'chroma_stft' were the only pitch-content features extracted; the rest all related to the timbre and texture of the tracks. All of the Audio features used are described as 'spectral features'.
Feature scaling was used in the creation of the following two datasets, in order that each feature played an equivalent role when put through the machine learning classifiers.

## 6. METHODOLOGY:

This section talks about the model used for the project. We used ANN (Artificial neural network )model.

**Artificial Neural Network**

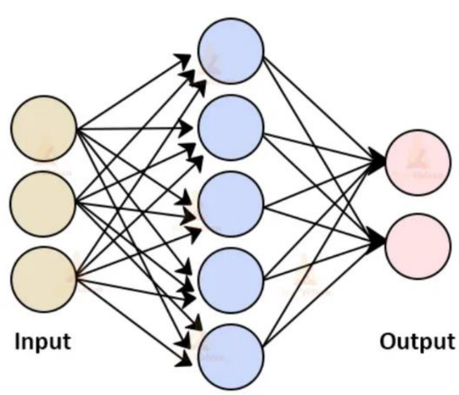
Artificial Neural Networks (ANN) are **multi-layer fully-connected neural nets** that look like the figure below. They consist of an input layer, multiple hidden layers, and an output layer. Every node in one layer is connected to every other node in the next layer.

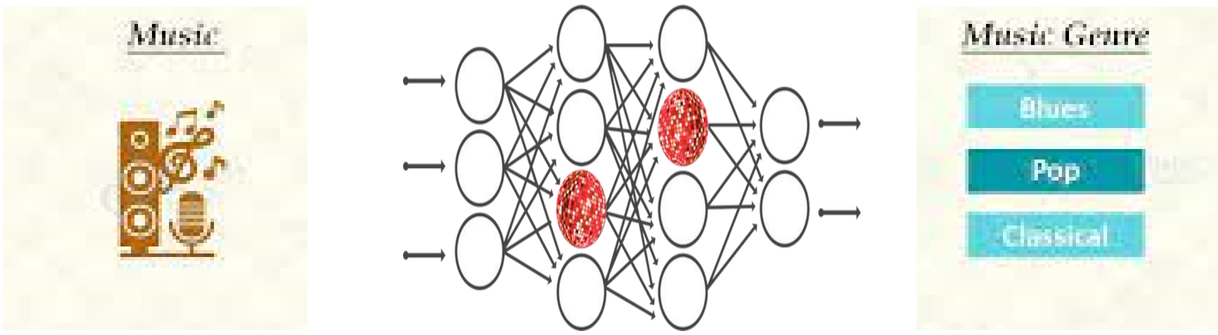
**Implementation of Artificial Neural Network(ANN)**

1. Import the Libraries.
2. Import the dataset.
3. Encoding the Categorical data.
4. Split the dataset for test and train.
5. Feature Scaling.
6. Import the Libraries.
7. Initialize our ANN model.
8. Adding the input layer and first hidden layer.

Artificial neural networks are a popular form of machine learning classifier, consisting of layers of connected nodes that transform input data to some kind of output. In the context of classification, neural networks learn models from training data, in order to predict the class of unseen data samples. Various kinds of neural network exist, including fully-connected, convolutional, and recurrent, and networks can be structured with varying numbers of nodes and layers according to design choices made by the engineer. A number of strategies exist for combating overfitting and reducing model complexity, and, when architected correctly, neural networks have proven to be especially effective for classification problems.

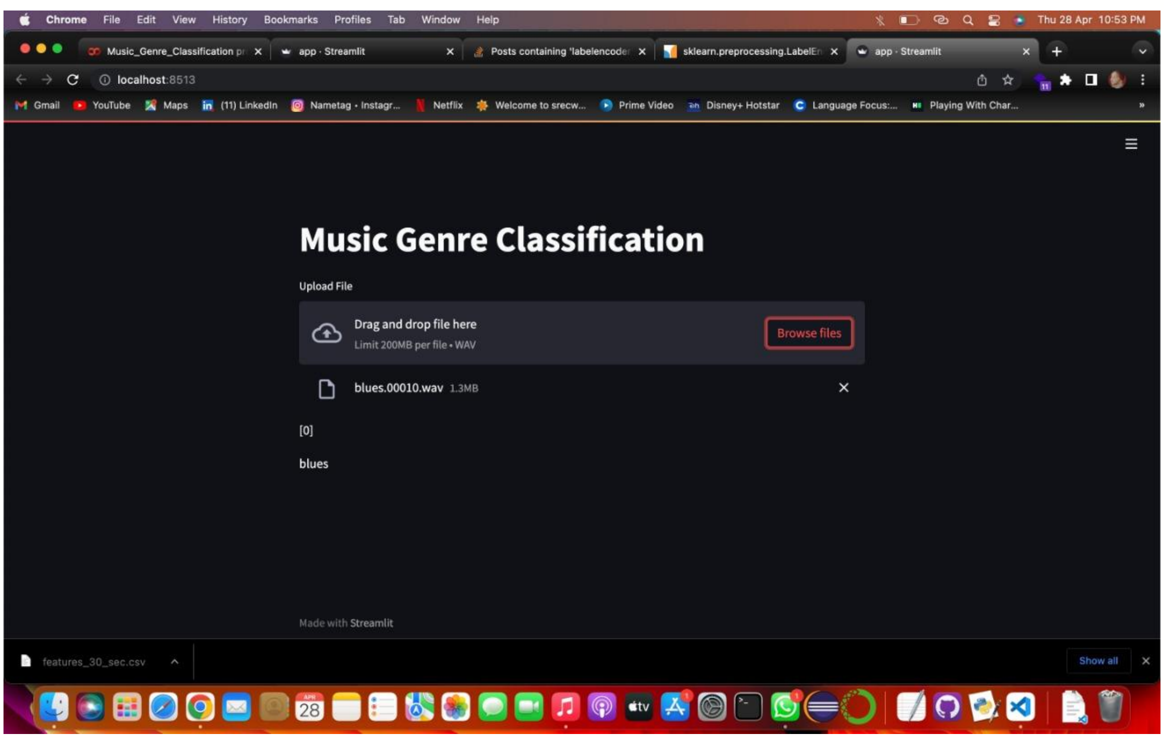
Neural Networks make use of loops and are able to 'remember' information between stages of the network's processing.

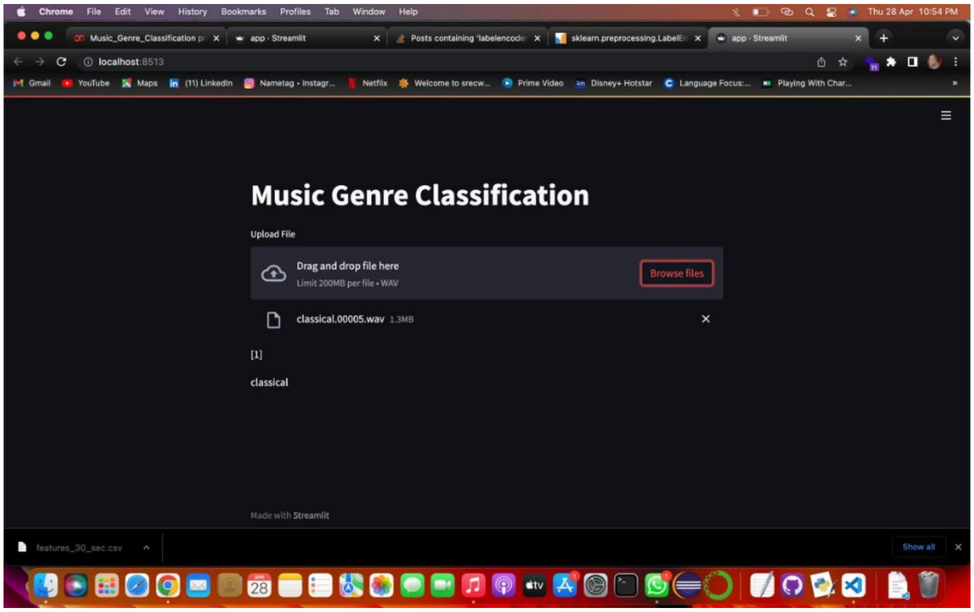


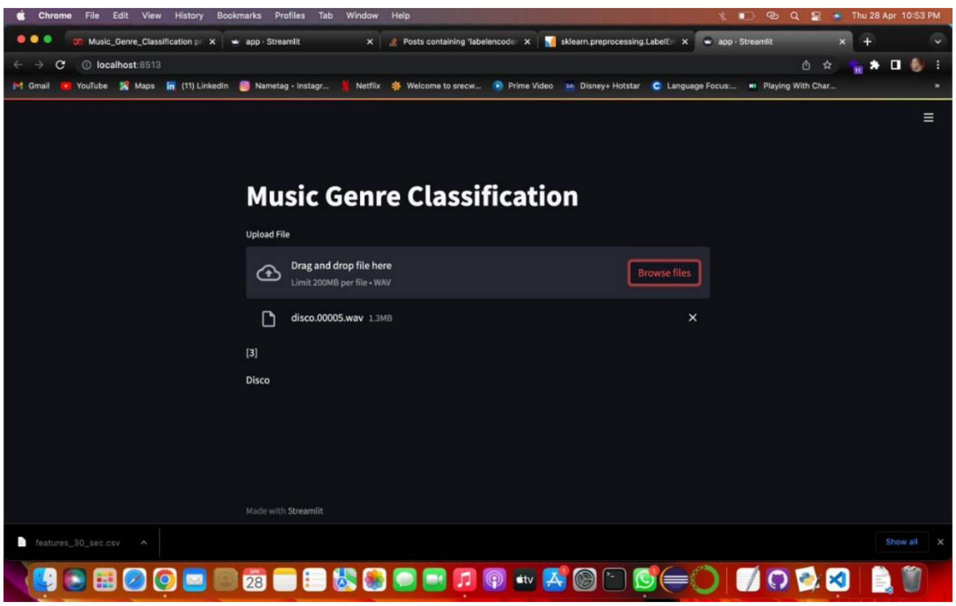



## 7. RESULTS:

### APPLICATION:






## 9. CONCLUSTION:

We saw how to develop a artificial neural network for music genre recognition. In this music genre classification project, we have developed a classifier on audio files to predict its genre. We work through this project on musicgenre classification data-set. It explains how to extract important features from audio files.

## 10. REFERENCES:

François Chollet et al. Keras

https://keras.io, 2015.

Mel Frequency Cepstral Coefficient (MFCC) tutorial

http://practicalcryptography.com/miscellaneous/machi ne-learning/guide-mel-frequency-cepstral-coefficientsmfccs/

Spectrogram computation in Signal Analyzer

https://www.mathworks.com/help/signal/ug/spectrog ram-computation-in-signal-analyzer.html

The dummy's guide to MFCC

https://medium.com/prathena/the-dummys-guide-tomfcc-aceab2450fd

Machine Learning GeeksforGeek

https://www.geeksforgeeks.org/machine-learning/

GTZAN Dataset Site :

http://marsyas.info/downloads/datasets.html

Musical Genre Classification with Convolutional Neural Networks by Leland Roberts :

https://towardsdatascience.com/musical-genreclassification-with-convolutionalneural-networksff04f9601a74