# Music Genre Classification using DNNs

## Introduction

- In our project, we aim to classify music based on its genre using feature extraction followed by machine learning classifiers.
- Our main focus is to study and analyze how different categories of audio signals consist of varying information in both the time and the frequency domain and thus exploit them to build features for accurate classification.

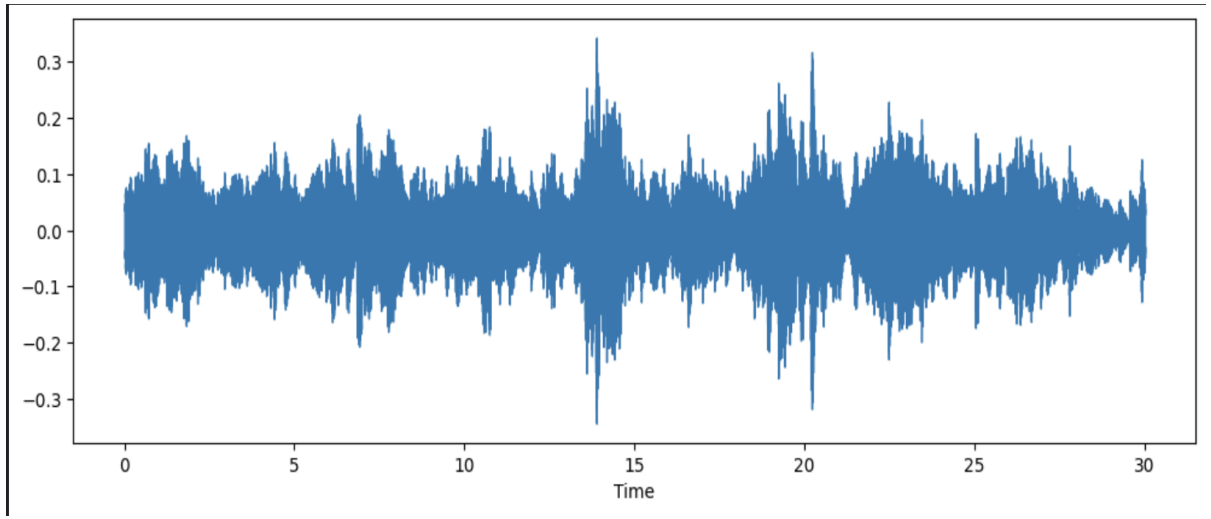
## Methodology

- Dataset acquisition and pre-processing
- Extraction of Features
- Building and Training a DNN Model on extracted features
- Classification

# Dataset Description

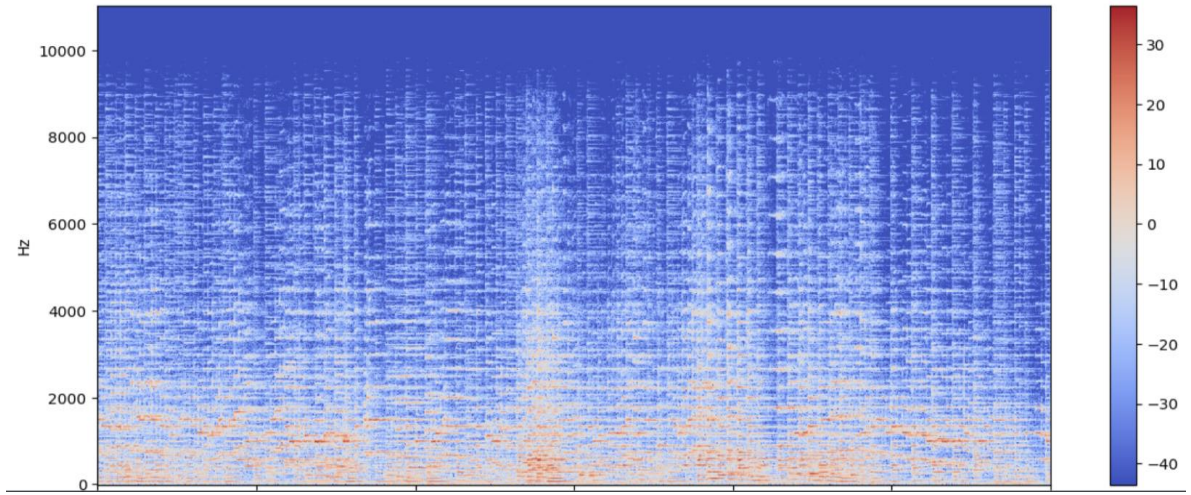
- We have used the GTZAN dataset. It contains 10 music genres, each genre has 100 audio clips in .wav format. Each audio clip is 30 seconds long, 22050 Hz Mono 16-bit file.
- **2 CSV files –** Containing features of the audio files.

## Visualization of an audio sample




Raw audio file of a Classical genre audio sample in Time domain

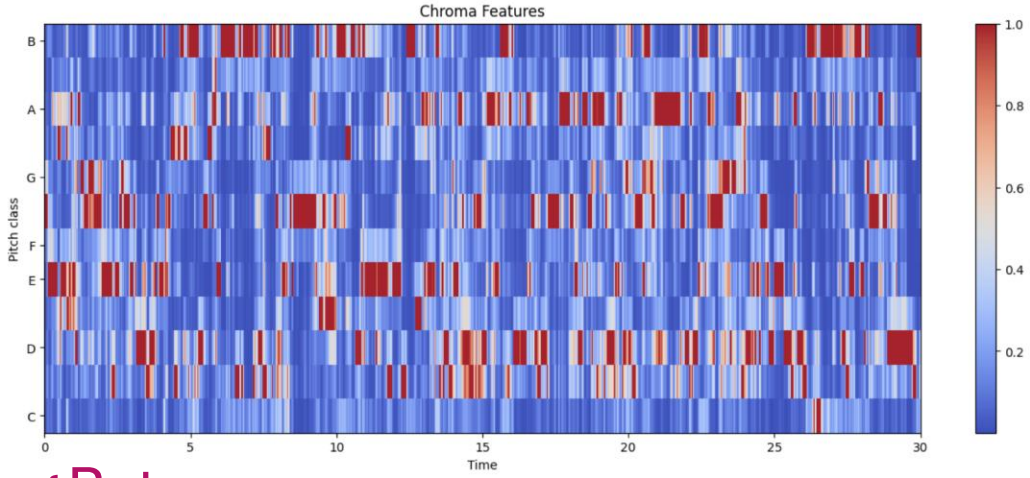
Spectrogram of same sample:
Visualization in Frequency domain

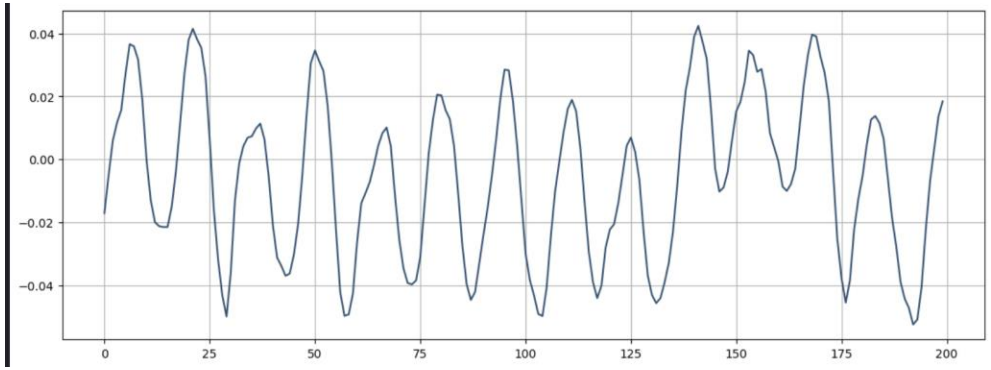


# Extraction of Features of audio sample

## Chroma Features

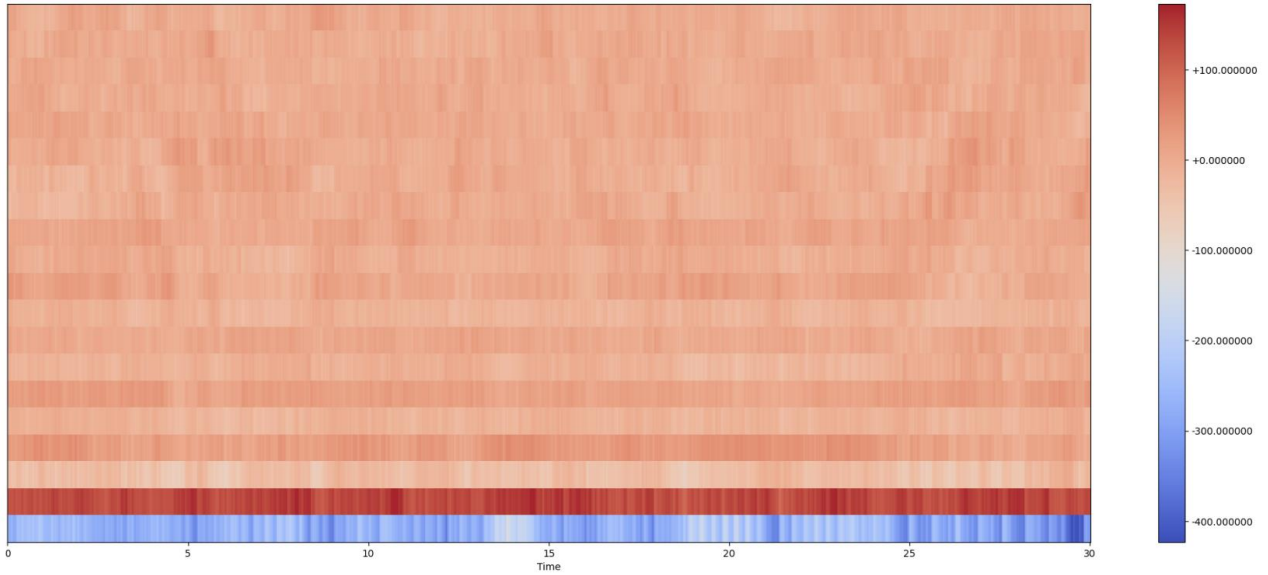
- It is a powerful tool for analyzing music features whose pitches can be meaningfully categorized and whose tuning approximates to the equal-tempered scale.




## Zero Crossing Rate

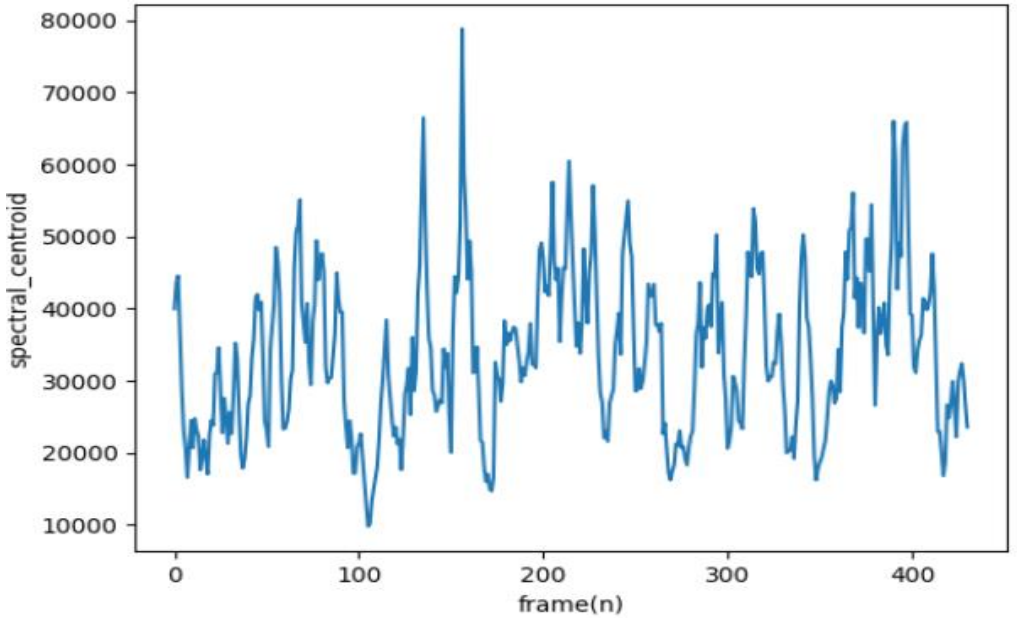
- Zero-crossing rate is a measure of the number of times in a given time interval/frame that the amplitude of the speech signals passes through a value of zero.
- The rate at which zero-crossings occur is a simple measure of the frequency content of a signal.




## Mel-Frequency Cepstral Coefficients

- MFCC represents a set of short term power spectrum characteristics of the sound and have been used in the state-of-the-art recognition and sound categorization techniques.

$$M(f) = 1125 \ln(1 + f/700)$$




## Spectral Centroid

- It describes where the "center of mass" for sound is.
- This feature describes the center of frequency at which most of the power in the signal is found.

$$SC = \frac{\sum_k kX(k)}{\sum_k X(k)}$$

## Spectral Roll off

- Spectral Rolloff is the frequency below which a specified 85%percentage of the total spectral energy lies.

$$\sum_{k<v} X(k) = (0.95)\sum_k X(k)$$

# Building a DNN model for Classification

## Model Description

- A DNN model with 4 dense layers and 1 FC layer with different no. of neurons per layer is implemented with Tensorflow framework.
- All of the hidden layers are using the ReLU activation function and the output layer uses the softmax function.
- Dropout is used to prevent overfitting.

## Model Training

- The described NN architecture model is trained on the above extracted features of input samples in GTZAN Dataset.
- We used the Adam optimizer for training the model. The epoch that was chosen for the training model is 500.
- The loss is calculated using the Sparse Categorical Cross entropy function.

## DNN Inference

- An inference pipeline to extract the above said features from the input sample is implemented.
- The output of the model is a vector of probability scores of genres with the highest probability score sample being the Genre of the input audio

## Results

- Test set accuracy is : 93.29%, Validation accuracy : 94.99% , with a Test loss : 0.529
- The model can classify the given audio into any one of 10 genres mentioned in the dataset.
- Our code can extract and visually depict the features of a given audio sample
- Kaggle Notebook Link

## Summary

- Our main focus was to examine how different generical audios differ in their spectral contents.
- So, we extracted the basic relevant spectral features for different audios and compared the differences first analyzing manually and then feeding the data into the machine learning classifier.
- In the end the features extracted were Chroma features, Zero Crossing Rate, MFCC, spectral centroid and spectral roll off.

## References


- http://ismir2001.ismir.net/pdf/tzanetakis.pdf
- https://farranaanjum05.medium.com/music-genre-classification-with-python-51bff77adfd6
- https://cse.iitk.ac.in/users/cs365/2015/_submissions/archit/report.pdf


# Thank You