Pattern Recognition Machine Learning

- Music Genre Classification

Member One
Deepak Bhatter
B22EE022

Member Two
Tushar Bhatt
B22CS056

Member Three
Rhythm Patni
B22CS043

Member Four
Rahul Sharma
B22EE051

Member Five
Prajjwal Dixit
B22ES002

Member Six
Mayank Agrawal
B22CS084

Problem Statement

Music plays a central role in our lives, and with the rise of streaming services, efficient music organisation becomes crucial. This project tackles music genre classification, aiming to develop a machine learning model that automatically identifies musical genres from audio data. This report details the chosen dataset, proposed approaches for feature extraction and classification.

Dataset

The GTZAN dataset is a collection of 1000 audio recordings, each 30 seconds long. The recordings are categorised into 10 different genres, with 100 files per genre and the genres are as follows:

1. Blues
2. Classical
3. Country
4. Disco
5. Hiphop
6. Jazz
7. Metal
8. Pop
9. Reggae
10. Rock

The audio files are stored in WAV format with a sampling rate of 22,050 Hz and 16-bit depth (mono channel).

The main data folder consists of two '.csv' files and two folders.

- The 'genres_original' has all the original audio files we need to analyse.The folder has 10 separate folders for each genre.
    - Each folder consists of 100 different audio clips of the same genre.
    - So, in total 1000 audio files are present as '.wav' in the genres_original folder.

- The 'images_original' has the images of the waveforms of all the songs we need to analyse. The folder has 10 separate folders for each genre.
    - Each of these folders consists of 100 images representing the waveforms of the corresponding audio clips in the 'genre_original' file.
    - So, the 'images_original' folder consists of 1000 waveforms as '.png' files.
- The main folder also contains 2 '.csv' files. One of which is the 'features_30.csv', and 'features_3.csv'. These files contain 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)

Proposed Approaches

We plan to perform the following tasks for our project:

- **Feature Extraction with PCA (Principal Component Analysis):** We will first extract relevant features from the audio data using Principal Component Analysis (PCA). PCA helps reduce dimensionality by identifying the most informative features in the data, which can improve model performance and training efficiency.

- **Classification Model Evaluation:** To establish a baseline and compare performance, we will train and evaluate several machine learning models on the extracted features. These models include K-Nearest Neighbors (KNN), Decision Trees, Support Vector Machines (SVM), Adaboost, and Logistic Regression. Evaluating their accuracy on a validation set will allow us to compare different approaches and find out which model performs the best.

- **Final Classification with Artificial Neural Networks (ANN):** We will implement an Artificial Neural Network (ANN) for the final music genre classification task. ANNs have proven effective in handling complex patterns in audio data, potentially leading to superior classification accuracy compared to the initial models.