Assignment 3

Problem Statement

To design and build CNN model for identification of Dysarthria disease

Data:

https://www.kaggle.com/datasets/iamhungundji/dysarthria-detection

Create a Kaggle notebook in this dataset. Perform all the tasks in this notebook.

Task:

1. Read data.csv in pandas
2. For any 1 random sample of dysarthric male, dysarthric female, non-dysarthric male, and non-dysarthric female, visualize
    a. Waveplot
    b. Spectrogram
    c. Zero Crossing rate
    d. Spectral Centroid and rolloff,
    e. MFCC
    f. Mel Spectrogram
3. Comment on the output.
4. Define a function to create 256 MFCC features for any given audio file.
5. Split the data into training and validation by a ratio of 90:10.
6. Build the CNN model with padding on a training set.
7. Evaluate the validation set.
8. Calculate and comment on ROC AUC score, recall score and confusion matrix.