

Assignment 5

Problem Statement

To build a model for different language word detection.

Data:

<https://www.kaggle.com/datasets/hazrat/urdu-speech-dataset>

Dataset contains a total of 2500 audio files of males and females, choose any 15 words out of available corpus (in PDF) for each available speaker. Please do mention the selected words in your solution notebook.

Task:

1. Prepare the dataframe containing selected words and audio file path.
2. Split the dataset into a training and validation set.
3. Apply MFCC on those audio files.
4. Build a model to detect the words from the speech using MFCC features.
5. Build the confusion matrix and F1-score on the validation set.
6. Record a sentence speaking atleast four words from the selected corpus, split the words from that sentence, test the model and provide your inference.
7. Publish the solution notebook in that Kaggle dataset.