Assignment 8

Problem Statement

Speech Analysis on European Languages

Data:

https://www.kaggle.com/datasets/bryanpark/spanish-single-speaker-speech-dataset
https://www.kaggle.com/datasets/bryanpark/french-single-speaker-speech-dataset
https://www.kaggle.com/datasets/bryanpark/german-single-speaker-speech-dataset
https://www.kaggle.com/datasets/bryanpark/dutch-single-speaker-speech-dataset

Choose anyone dataset

Task:

1. Read at least ~2000 audio files and transcripts from the data.
2. Calculate the MFCC components for the selected audio files.
3. Prepare the input data for the model such that, given any transcript, predict the next word with the help of the MFCC features.
4. For MFCC, build the CNN model and for transcript build any sequence model.
5. Prepare the preprocessing pipeline to predict the transcript for any raw audio files. Evaluate the model for at least 10-12 audio files.
6. Publish your work in the Kaggle.