Speech Processing Lab Assignment – 2

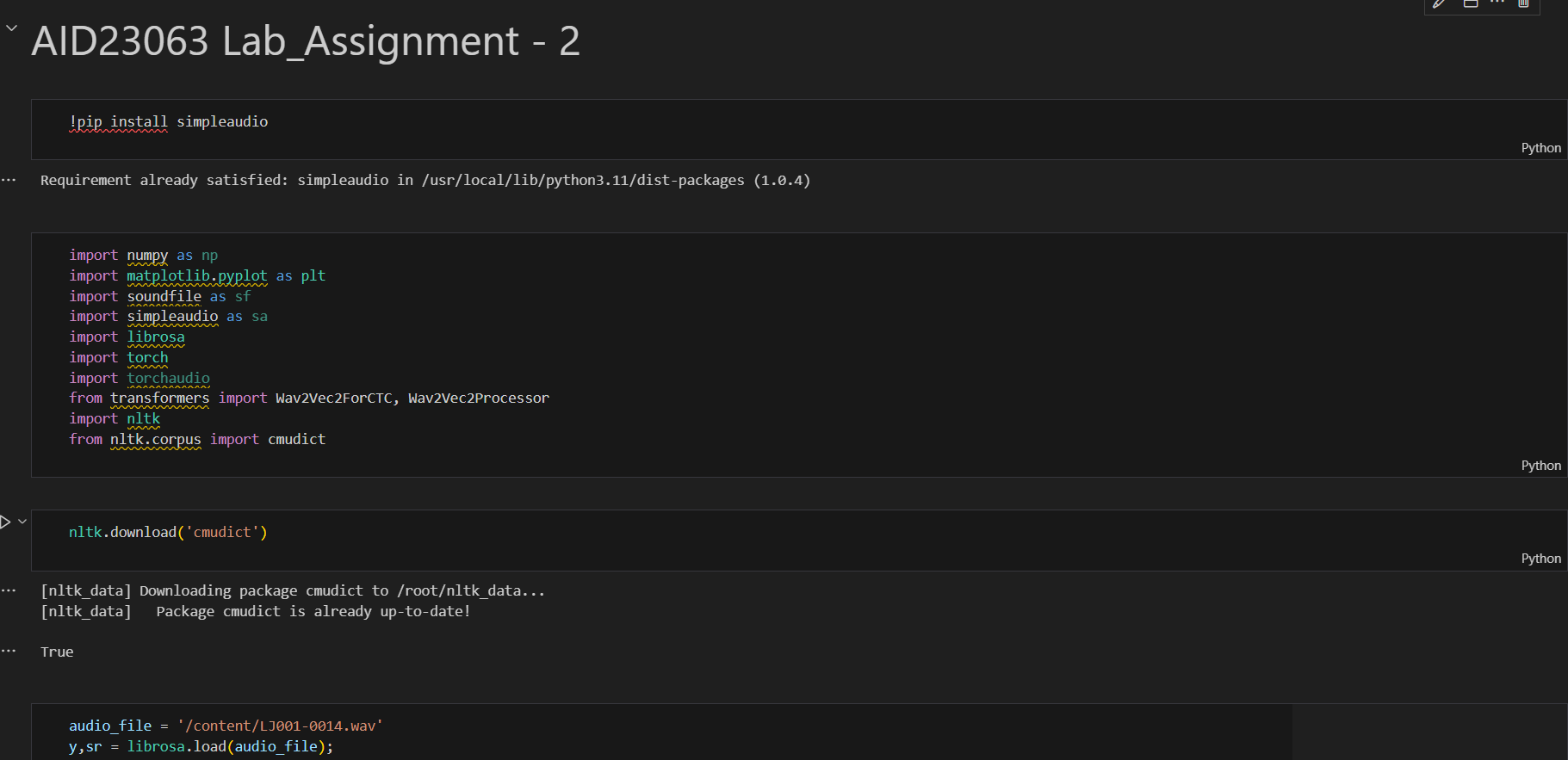
Name: Paruchuri Sai RegNo: BL.EN.U4AID23063

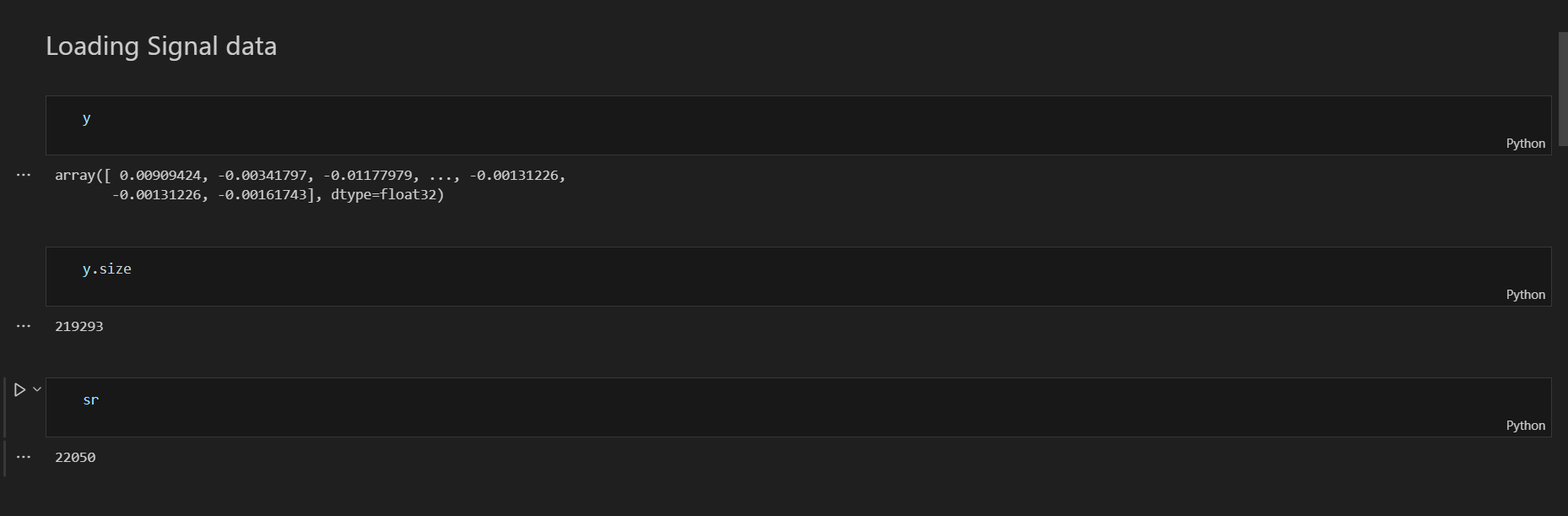
Phoneme Extraction and Classification from Speech Signals

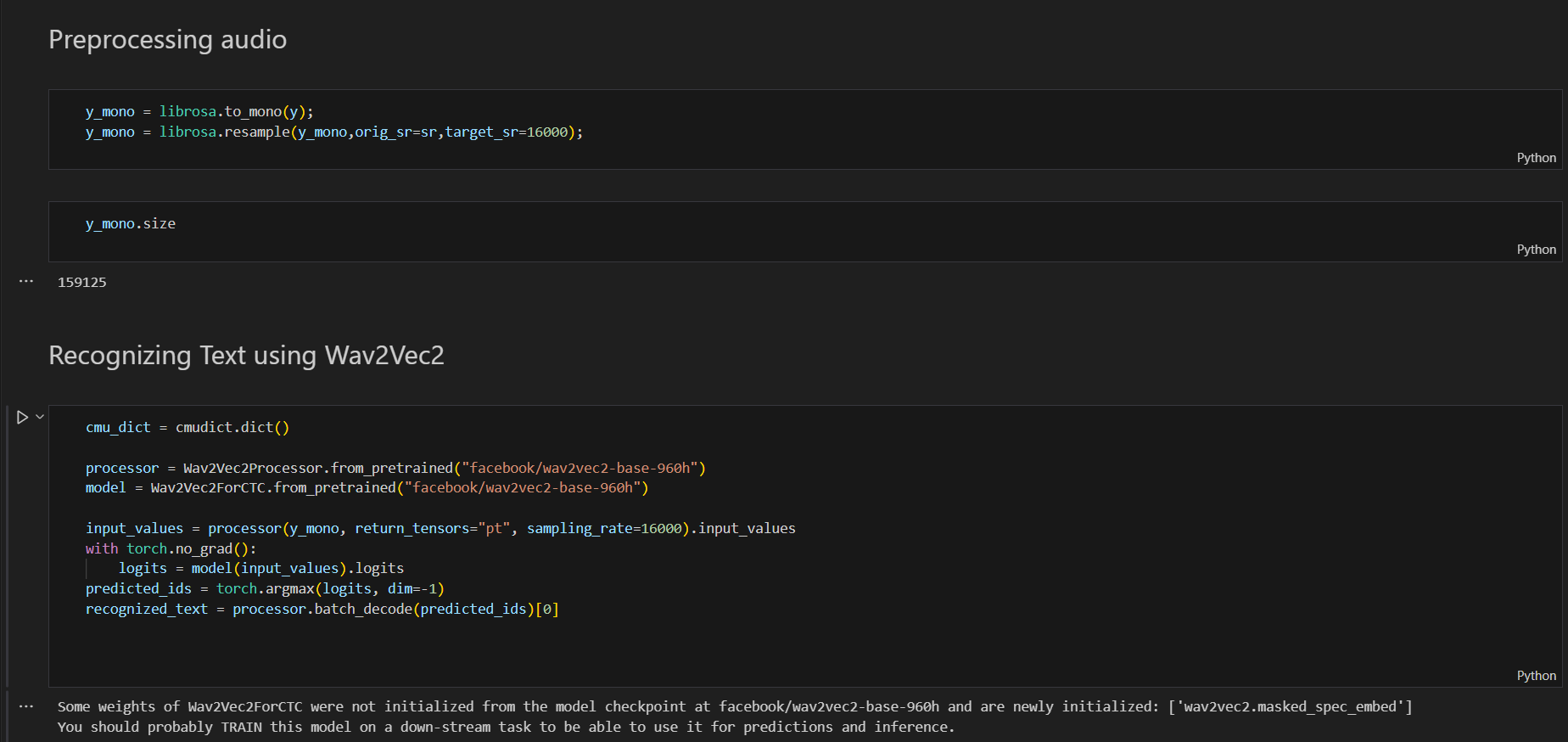
Dataset Description: The speech file is taken from LJ Speech Dataset from Kaggle website which contains short audio clips of speakers reading passages from books. The speech has 9 seconds duration.

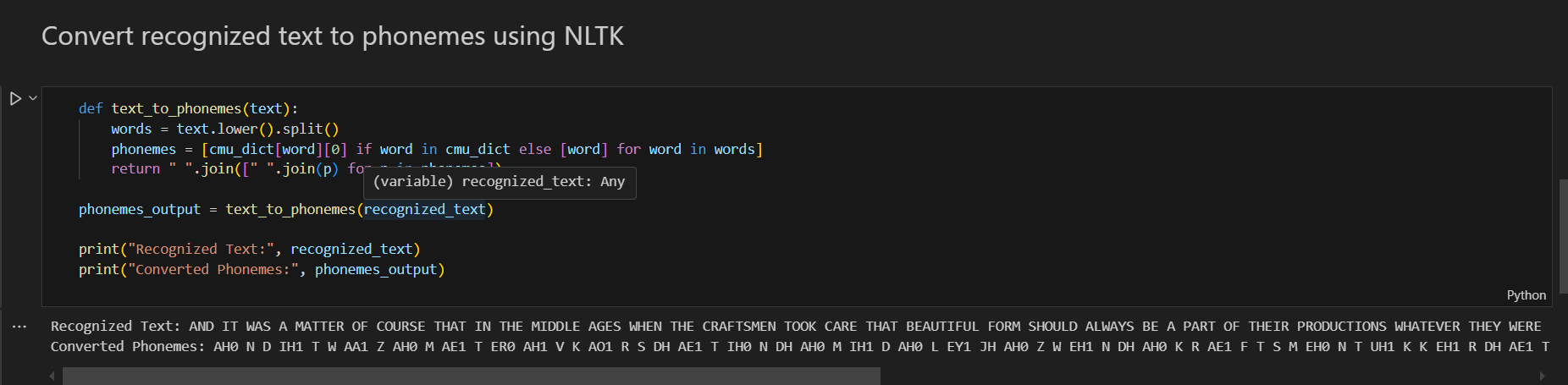
Objective: The objective of this assignment is to process a speech signal, extract specific phonemes, and visualize their waveforms while labelling them.

Code and Output:









Recognized Text: AND IT WAS A MATTER OF COURSE THAT IN THE MIDDLE AGES WHEN THE CRAFTSMEN TOOK CARE THAT BEAUTIFUL FORM SHOULD ALWAYS BE A PART OF THEIR PRODUCTIONS WHATEVER THEY WERE

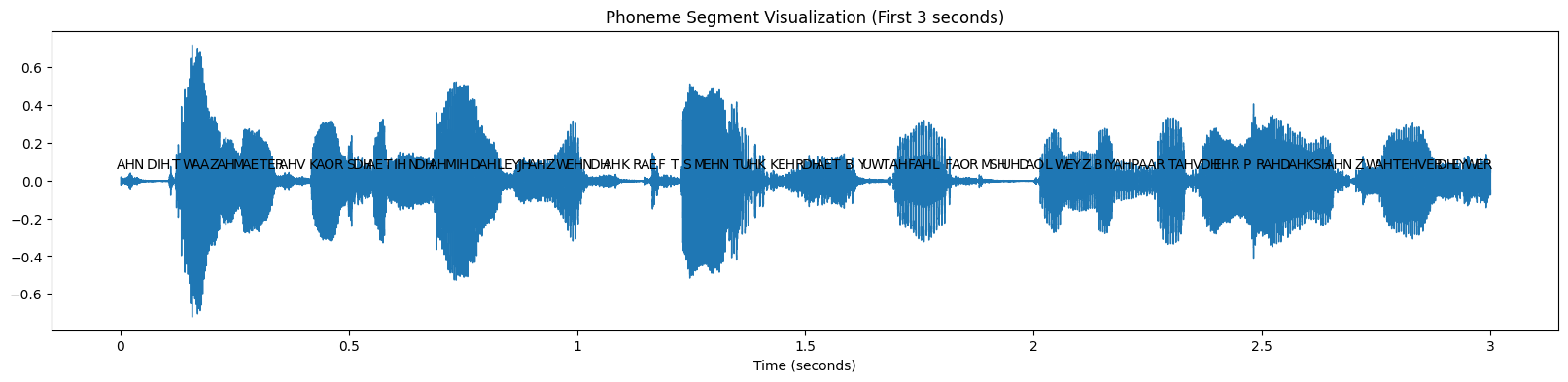
Converted Phonemes: AH0 N D IH1 T W AA1 Z AH0 M AE1 T ER0 AH1 V K AO1 R S DH AE1 T IH0 N DH AH0 M IH1 D AH0 L EY1 JH AH0 Z W EH1 N DH AH0 K R AE1 F T S M EH0 N T UH1 K K EH1 R DH AE1 T B Y UW1 T AH0 F AH0 L F AO1 R M SH UH1 D AO1 L W EY2 Z B IY1 AH0 P AA1 R T AH1 V DH EH1 R P R AH0 D AH1 K SH AH0 N Z W AH2 T EH1 V ER0 DH EY1 W ER0



AH N D IH T W AA Z AH M AE T ER AH V K AO R S DH AE T IH N DH AH M IH D AH L EY JH AH Z W EH N DH AH K R AE F T S M EH N T UH K K EH R DH AE T B Y UW T AH F AH L F AO R M SH UH D AO L W EY Z B IY AH P AA R T AH V DH EH R P R AH D AH K SH AH N Z W AH T EH V ER DH EY W ER







Conclusion:

We have seen in this assignment how to process a speech signal, extract specific phonemes, and visualize their waveforms while labeling and plotting them. we have converted audio signal into 16khz sampling rate. Then used Wav2Vec2 to recognize phonemes and then extracted a phoneme segment on a time interval of first 3 seconds and then visualized extracted phoneme by matching it with recognized phonemes.

GitHub code Link:

https://github.com/SaiParuchuri123/Speech\_Processing