**In this link:**

* I have uploaded all the three videos for the three audios provided from your side.
* Also the excel sheets of the grammatical mistakes for the provided three audios.
* Also the screenshots for the times where the model has detected “aaa...” or “uhh...” or “umm...”.

In the video and screenshots , you can see a table where there is a start time and an end time for the particular pronunciation.

The format of the table is like-

start\_time | end\_time | type | content | confidence

Along with this is have provided the audio file where I have recorded my audio with wrong pronunciations and I have provided a .txt file with the transcript of the corrected pronunciation done by the model.