Reporting No:3 Week No:3 From :26/05/2025 To :30/05/2025

College ID: 22IT121

Project Title: Speech-to-Phoneme Normalization with CMU Dictionary and TTS

WEEKLY REPORT

Work done in last week (Attach supporting Documents):

26/05/2025 Monday

9:30 – 11:30	Explored methods to extract **phonemes for whole sentences**, not just the first word.
12:00– 2:30	Integrated this functionality into the Flask pipeline.Tested with different sentences and phrases.
3:00– 5:00	Checked for fallback when a word is not in CMU and implemented a fallback to spelling the word.

27/05/2025 Tuesday

9:30 – 11:30	Added a **simple logging feature** to track transcription and phoneme extraction in a CSV file.
12:00– 2:30	Integrated logging into the UI to view previous recordings.
3:00– 5:00	Tested with multiple phrases to validate robustness and accuracy.

28/05/2025 Wednesday

9:30 – 11:30	Explored **deployment options** for Flask application.
12:00– 2:30	Started preparing files and requirements (`requirements.txt`) for deployment.
3:00– 5:00	Tested deployment on **PythonAnywhere**, addressed issues with audio recording and directory structure.

29/05/2025 Thursday

9:30 – 11:30	Integrated **WebRTC** to allow **direct browser recording** instead of `sounddevice`.
12:00– 2:30	Checked compatibility across different browsers (Chrome, Edge, Firefox).
3:00– 5:00	Added fallback messages for unsupported browsers and cleaned up UI messages.

30/05/2025 Friday

9:30 – 11:30	Final testing of the **deployed application**.Checked for latency, transcription accuracy, and phoneme delivery.
12:00– 2:30	Prepared final documents and a short video walkthrough of the application.
3:00– 5:00	Backup of files, code, and reports for submission and future improvement.

Reason for incomplete work:

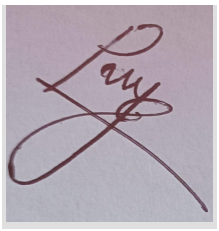
1. N/A

Plans for next week:

1. The Flask application lets users:
2. Record directly from their browser's microphone.
3. Transcribe their speech into text.
4. Retrieve CMU phonemes for the first word or the whole sentence.
5. Hear back the pronounced word with US English TTS.
6. View a log of previous recordings and transcriptions.
7. Handle fallback gracefully when a word isn't in the CMU dictionary.
8. Prepare for future expansion to scoring, phone apps, or richer ui.

References:

1. Speech-to-Text: SpeechRecognition, Google API
2. Text-to-Speech: pyttsx3
3. Phoneme Dictionary: NLTK cmudict
4. Audio: Sounddevice
5. Deployment: PythonAnywhere, WebRTC
6. UI: Flask, HTML, CSS



Signature of External Guide **Signature of Internal Guide**

Student Id: 22IT121
Student Name: VYOM MUKESHKUMAR PATEL