Reporting No:2 Week No:2 From :19/05/2025 To :23/05/2025

College ID: 22IT121

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

WEEKLY REPORT

Work done in last week (Attach supporting Documents):

19/05/2025 Monday

9:30 – 11:30	Reviewed existing CLI-based speech-to-phoneme system and planned web-based UI flow using Flask.
12:00– 2:30	Initialized Flask project structure with necessary folders for templates and static files. Installed Flask and verified server run.`avfile` to save spoken input as `.wav` files.
3:00– 5:00	Implemented basic Flask route to handle GET and POST requests and connected it to the transcription pipeline.

20/05/2025 Tuesday

9:30 – 11:30	Integrated audio recording functionality using `sounddevice` directly within the Flask POST handler.
12:00– 2:30	Tested audio saving mechanism to `static/input.wav` and ensured compatibility with SpeechRecognition and scipy.

3:00– 5:00	Created the transcription route to convert recorded audio into text and handle exceptions gracefully.

21/05/2025 Wednesday

9:30 – 11:30	Extracted first word from transcription result and looked up phonemes from CMU Pronouncing Dictionary using NLTK.
12:00– 2:30	Implemented `pyttsx3` speech synthesis for the recognized word with filtering of US English voices.
3:00– 5:00	Tested the pipeline from recording → transcription → phoneme lookup → speech synthesis.

22/05/2025 Thursday

9:30 – 11:30	Designed and implemented a responsive front-end using HTML and embedded Flask variables to display transcription and phonemes.
12:00– 2:30	Connected web interface button to the backend pipeline; verified end-to-end data flow from user interaction to audio response.
3:00– 5:00	Applied basic CSS styles to improve UI readability and tested the application across different browsers.

23/05/2025 Friday

9:30 – 11:30	Handled edge cases such as unclear audio and words not present in the CMU dictionary, with appropriate UI feedback.
12:00– 2:30	Tested full web application across multiple inputs and verified accuracy of phoneme extraction.
3:00– 5:00	Prepared final version of the Flask UI project for submission and documentation.

Reason for incomplete work:

1. N/A

Plans for next week:

1. Add camera-based audio capture (mobile mic/browser mic).
2. Enhance GUI with phoneme playback and IPA format support.
3. Deploy Flask application on PythonAnywhere or Render for public demo.
4.

References:

1. Flask Official Documentation
2. SpeechRecognition and Pyttsx3 Python libraries

3. NLTK and CMU Pronouncing Dictionary usage guides
4. HTML/CSS templates from W3Schools and FlaskJinja resources

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

Student Id: 22IT121
Student Name: VYOM MUKESHKUMAR PATEL