

MAI572 Speech Processing & Recognition

Department of Computer Science, Christ University Central Campus

Lab Exercise III

Speech/Audio to Text Conversion

You have been hired as an AI engineer by a tech startup that focuses on enhancing accessibility for people with disabilities. One of your primary tasks is to develop a feature that allows users to control devices and input text via voice commands.

The first version of this system requires you to implement a speech-to-text application that takes spoken commands and converts them into text in real time. This application will serve as the foundation for future development, where you will integrate it with various smart devices and software. Your tasks are listed below:

(1) Tasks

- (a) Record spoken input using the microphone.
(Alternatively, you may use any speech signal from a dataset)
- (b) Convert the speech into text.
- (c) Display the converted text on the screen.
- (d) Handle scenarios where speech is unclear or the system cannot connect to the recognition service.

(Hint: Exception handling)

- (i) Unclear Speech (speaking indistinctly, e.g., mumbling,)
- (e) Provide meaningful feedback to the user at various stages and situations.

For example,

- (i) 'Speak something'
- (ii) 'Recognizing'
- (iii) 'Speech recognized: "Turn on the lights in the living room.'

- (iv) 'Speech successfully converted to text!'
- (v) 'Speech Recognition could not understand audio, Please try speaking more clearly'
- (f) Write a brief note/inference about the execution of the total system.

Evaluation Rubrics:-

- (1) Execution/Implementation: 4 marks.
- (2) Concept clarity: 2 marks
- (3) Complexity/self-learning: 2 marks.
- (4) Documentation & Writing the inference: 2 marks

Submission Guidelines:-

- Generate the single .pdf file for the given questions separately.
- Upload the pdf files in Google Classroom on or before the deadline mentioned.