TECH SAKSHAM

FINAL PROJECT REPORT

SPEECH RECOGNITION

MIT COLLEGE OF ARTS AND SCIENCE FOR WOMEN, MUSIRI

ROLL. NO.	NAME
CB20S202963	Priyadharsini D
CB20S202977	Tamilselvi R
CB20S202980	Yuvasri M
CB20S202973	Sneka P

Trainer name: Mayank Shrivastava

Master name: Mayank Shrivastava

ABSTRACT

Speech recognition technology has seen a rapid evolution over the years and has become an essential tool in various industries, such as healthcare, finance, and customer service. In this project, we aim to develop a text to speech and speech to text application using Python and a graphical user interface (GUI). The application provides a user-friendly interface for users to convert text to speech and vice versa by leveraging deep learning algorithms. The project architecture includes two main components: the text to speech component and the speech to text component. The text to speech component uses the Google Text-to-Speech (gTTS) API to convert text to audio, and the speech to text component utilizes the SpeechRecognition library and the Google Cloud Speech API to recognize speech and convert it to text. The application's GUI is developed using the Tkinter library, and the application's functionalities are integrated with various event handlers. The report provides an overview of the project's architecture, its working details, and the tools and services required for its development. We conclude that the application can be useful in various fields, including education, entertainment, and communication, and has the potential to be extended to more advanced speech recognition and synthesis applications.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Speech recognition is an emerging technology that has become increasingly popular due to its ability to recognize human speech and convert it into digital format. This technology is used in various applications such as voice assistants, dictation software, and automated call centers. In this project, we have implemented speech recognition technology in Python using the SpeechRecognition library and Text-to-Speech conversion using the gTTS library. We have also created a GUI for our application using Tkinter.

CHAPTER 2

SERVICES AND TOOLS REQUIRED:

- Python
- SpeechRecognition library
- gTTS library
- Tkinter library
- Playsound library.

CHAPTER 3

PROJECT ARCHITECTURE:

The project architecture is divided into two parts - Speech to Text and Text to Speech. The Speech to Text module uses the SpeechRecognition library to convert speech into text, while the Text

to Speech module uses the gTTS library to convert text into speech. The GUI is created using the Tkinter library.

Fig. Text to speech

Fig. Speech to Text

CHAPTER 4

ARCHITECTURE WORKING DETAILS

4.1 Speech to Text:

The Speech to Text module uses the SpeechRecognition library to convert speech into text. We start by initializing a recognizer object and a microphone object. We then use the adjust_for_ambient_noise() method to cancel out any ambient noise and then use the listen() method to capture the user's speech. The captured speech is then sent to the

Google Speech Recognition API using the recognize_google() method, which returns the recognized text.

4.2 Text to Speech:

The Text to Speech module uses the gTTS library to convert text into speech. We start by initializing a gTTS object and setting the language and speed of speech. We then use the save() method to save the speech as an MP3 file and the playsound() method to play the speech.

4.3 GUI:

The GUI is created using the Tkinter library. We create a main window and add two frames - one for Speech to Text and the other for Text to Speech. We add labels and entry boxes for input and output and buttons for performing the conversion. We also add a button for listening to the user's speech.

CHAPTER 5

CONCLUSION

In this project, we have successfully implemented speech recognition and text to speech conversion using Python and created a GUI using Tkinter. The application has various use cases such as in voice assistants, dictation software, and automated call centers.

CHAPTER 6

REFERENCES

- SpeechRecognition Library Documentation: **https://pypi.org/project/SpeechRecognition/**
- gTTS Library Documentation: **https://pypi.org/project/gTTS/**
- Tkinter Library Documentation: **https://docs.python.org/3/library/tkinter.html**
- Playsound Library Documentation: **https://pypi.org/project/playsound/**

CODE

Github link for project code
https://github.com/DSPriyadharsini/SpeechRecognition/blob/main/main()