IDC409 Project

Group-3:- Krishna Aggarwal (MS21169)

Kanika (MS21035)

Sakshi A.Ninawe (MS21023)

Title: - **Speech-to-text Recognition**

Speech-to-text technology, also known as automatic speech recognition (ASR), is a technology that converts spoken language into written text. This technology has various applications and canbe used in a wide range of fields, including transcription services, voice assistants, accessibility tools, and more.

The Code Analysis: -

- **Recording Voice**

- **Playing Audio**

- **Converting to text**

Packages used: -

- **Sounddevice: -** Sounddevice is a Python module that provides bindings for the PortAudio library, which is a cross-platform C

library for real-time audio input and output. It makes it easy to record the input sound into a NumPy array**.**

- **Wavio: -** Wavio is a python module for reading and writing WAV files using NumPy arrays. It provides a simple and convenient way to manipulate audio data and we used this module to record the file from data in NumPy array.

- **Winsound: -** In Jupyter Notebook or JupyterLab, we can use the winsound module to play simple sound alerts on Windows. The winsound module provides a way to interact with the Windows Sound API.

  The winsound module is specific to Windows, and it may not workon other operating systems

- **AssemblyAI: -** AssemblyAI is a powerful speech-to-text API that allows you to transcribe audio and video files into text. It offers a variety of features, including the ability to transcribe multiple languages, identify speakers, and generate summaries. AssemblyAI provides a Python SDK that makes it easy to integrate its transcription capabilities into your Python applications. The SDK is well-documented and provides a simple API for transcribing audio and video files.