**University Living Campus Hiring Assignment**

Vishnu Pratap Singh (2022PAI7312)

M.Tech. (CSE-Artificial Intelligence)

NSUT, Delhi

**Task :-** (Task 2nd ) : Voice-to-Text System Development.

**Objective:-** The primary objective of this project is to design and implement a robust voice-to-text (speech recognition) system capable of accurately transcribing spoken language into textual form. The system should be able to process audio input from various sources and environments, converting it into written text with high accuracy and efficiency.

**Solution Approach:-**

**Step 1.** Make sure you have installed the required libraries first.

PyAudio, Numpy, and more requirements.

whisper install pip Installing

Pyaudio using pip

pip install Numpy

**Step 2.** Install Whisper: Whisper is an automated speech recognition (ASR) system that was developed using supervised multilingual and multitask online data. simply utilizing the Terminal command "pip install whisper." and add Whisper to the notebook Jupyter.

**Step 3.** Load the Whisper base model (because it is appropriate for the task at hand).

**Step 4**. We'll grant access to the system microphone so that real-time audio capture from the microphone may be done using the PyAudio module.

**Step 5.** The whisper base model only recognizes audio files up to 30 seconds, thus trim the audio to 30 seconds.

The parameters are as follows:

* Use Pyaudio to set the audio format to 16-bit PCM.(paInt16)
* Select 1 (mono) for the number of audio channels.
* A sampling rate of 16000 Hz should be set.
* Assign 1024 as the buffer size (chunk size).

**Step 6.** The Whisper model will be used to process and transcribe the audio frames.

**Step 7.** The assistance of the Command Code, The Whisiper Model is now able to record audio from the System Microphone; pressing the Key interrupt will end the recording. And it begins to generate a transcript after recording.

**Step 8**. Verify the outcome and the accuracy of the provided outcome (Transcript).

**Input:-** User-controlled real-time audio recording.

**Output:-** Text (Transcript based on Realtime Audio Input)

**Software Used:-** Platform – Jupyter Notebook (for python code)

Model - Whisper by Open AI (pre train multilingual model)