**TensorGo Assesment**

Mohammad Kaif (2022PAI7313)

M.Tech. (CSE-AI&ML)

NSUT- Delhi

**Task :-** Building a Multilingual Speech Recognition Model for RAG Without Training.

**Topic:-** Computer Vision and NLP.

**Solution Approach:-**

**Step 1-** install whisper (Whisper is an automatic speech recognition (ASR) system trained on multilingual and multitask supervised data collected from the web).by using Command “pip install whisper” on Terminal.

**Step 2-** import Whisper to the jupyter notebook.

**Step 3-** Load the base Model of whisper (as it is Suitable for Given Task)**.**

**Step 4**- import the audio file for transcription and add file path to the jupyter notebook.

**Step 5**- Trim the audio to 30 second time as Whisper base model only recognize audio file up to 30 second.

**Step 6**- Decode , Translate command to the audio file in notebook.

**Step 7**- Now with the help Command Code, Whsiper Model generate the transcript for the Loaded audio file successfully.

**Step 8**- Check the Result and Accuracy of the given result (Transcript).

**Input:-** An Audio file contain dialogue from the Movie “The Dark Knight” time span of 20 second. In English Language and Accent is American.

**Output:-** “Now, our operation is small, but there's a lot of potential for aggressive expansion. So which of you fine gentlemen would like to join our team?

Oh, there's only one spot open right now, so we're gonna have tryouts.”

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

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

Data - Sample Audio File.