

What Problem I am trying to solve

In short, it's a multilingual voice assistant, that can help

1. Not only to reduce the language barrier in using cutting edge technologies but tries to make everyday life a bit easier.
2. It also increases accessibility of the technology that can be helpful for people who have some kind of disability.

With the use of ChatGPT & Whisper API from Open API and by using Google Translate library from Python and Text to Speech API from Google Cloud I have created a web app that can take voice input from user and perform multiple tasks based on User preference, such as language Translation and/or Communication with chatGPT to get information.

How does the App Works

- It takes a voice input from user and user chooses an option whether he/she just want to translate the given sentence or want to get information from chatgpt.
- The app then translates the given audio to English and also remembers what was the original language of the audio. It is done using Whisper API.
- If the preference was only to translate, then using chat completion the app translates it to that desired language.
- Then convert the text to audio using using Text-To-Speech API from Google Cloud.
- At last it gives out the out both as Text(In English) and Audio output which is in the asked language.
- But If the user wants to interact with chatGPT then, the app first translate and transcribe the audio in English using Whisper API and sends the message to Chatgpt.
- After getting the response back from chatGPT, the app then translates the message to the original language, in which the audio was taken, using Google Translate Python library
- At least it convert that text to speech using TTS API from Google Cloud and provides a voice output(in the same language the voice input was). App also shows the output in English text as well.

Who can Use this

The use of this app is not confined in any particular space or domain. It can be used in everyday life and for business purposes as well.

With these capabilities, a person from any domain can use this app to get knowledge out of it, for example travel, any kind of business etc. For example:

1. a person from Europe wants to travel a city in India he can use the app to not only trip his/her plan in that city but also can communicate in the language he/she wants, obviously there would some cases where the meaning might get changed during translation but that part is dependent on the accuracy of the API.
2. Same for Business or Consulting with doctor.

Broadly It will be able to solve the following problem:

1. Reduce the language barrier. Like using a service that accepts query in only a certain languages. For example, using this app a person can ask question in Bengali, Hindi to chatGPT and get answer back in the same language.
2. Handsfree use of the services and it will also increase the accessibility of these services, for people who have some kind of impairment.
3. The language translation part the application can be used to learn new languages and can also be used as a utility to understand signs written in a language that the user don't know.
4. Also real time translation will be a great help in terms of communication between people who doesn't have a common language to speak in.

Future Improvement

- Right now the app only supports Translation to Bengali, Hindi and English. In future it will support more languages.
- Also, the processing can be more frugal. A device without GPU will take more time process these requests. To resolve this we can use smaller models for Whisper
- Also processing of the inputs can be handled in more resource restrictive way depending upon user voice inputs.