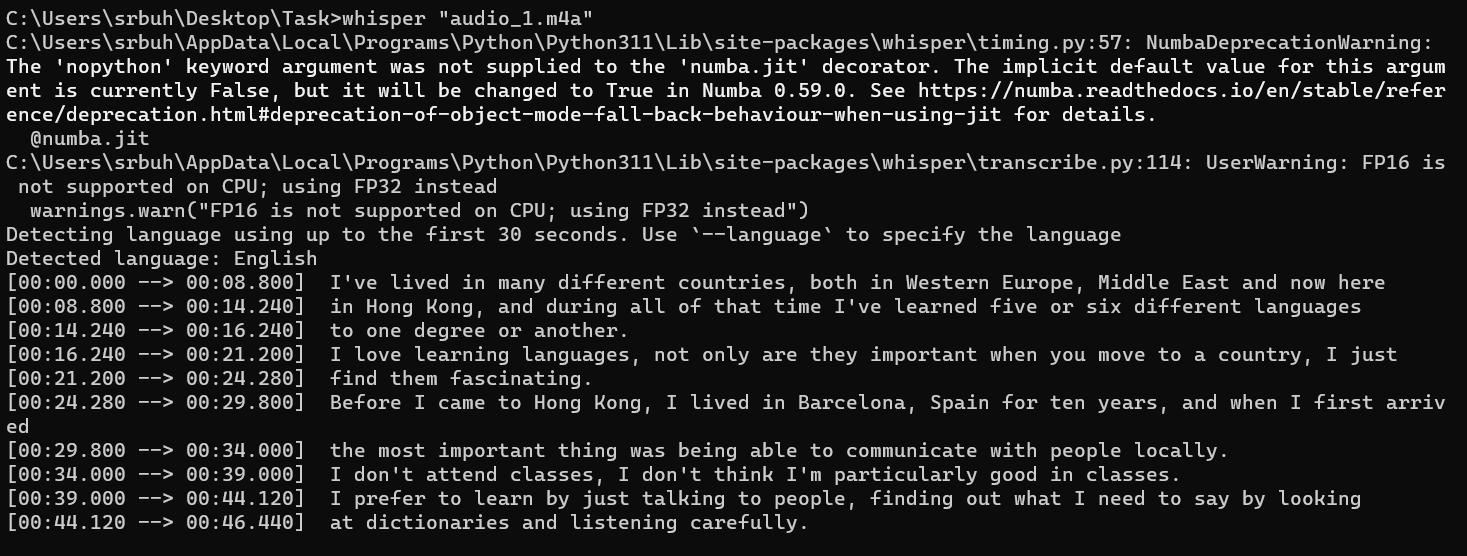
First I explored the whisper model, its structure and how it works. Then I tried to find some dataset of audios, but in the process I decided that I could record myself, and why not, I could added some noise too. For each audio, I created a text document (rec\_i , i=1,2,3,4) in which I put the correct text for that recording.

I did the necessary installs for whisper and then called it for all audios.



It gave me several files (such as .txt, .json, .srt , etc.) containing the text read by whispel.

Then I moved on to the next part of the task։ WER. I understood what it is, how the distance is determined․ I came across a code showing how to put WER between sentences. I tried it on my docs and the result didn't seem too bad 😊 .