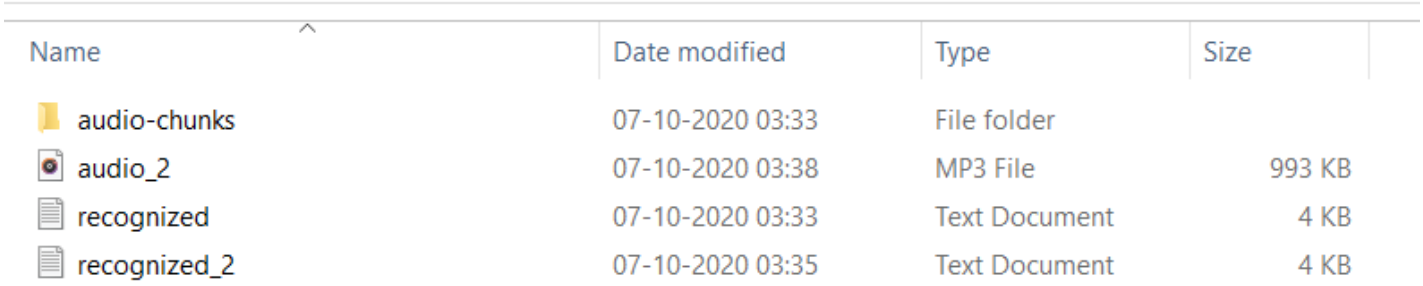
Project Name: Speech-Text All Large-Files Conversions

Table of Content

Demo



Name	Date modified	Type	Size
audio-chunks	07-10-2020 03:33	File folder	
audio_2	07-10-2020 03:38	MP3 File	993 KB
recognized	07-10-2020 03:33	Text Document	4 KB
recognized_2	07-10-2020 03:35	Text Document	4 KB

Overview

This is a Speech to Text and Text to Speech Converter for performing paragraphs or big chunk of text.
This repository contains the code for Speech Text Conversion using python's various libraries.
It used SpeechRecognition, googletrans, gTTS libraries.
These libraries help to perform individually one particular transformation.
Using SpeechRecognition, converted speech/voice to text format.
Using googletrans, converted output text paragraph to specified language.
Through gTTS, converted above translated text document to speech by giving voice to it.
Here I have used speechrecoginition, is to convert a pre-recorded long audio to text.
The purpose of creating this repository is to try all possible type of conversion with paragraph structure.
These python libraries raised knowledge in discovering these libraries with practical use of it.
It leads to grow in my AI repository.
The above screenshot will help you in understanding the flow of output.

Motivation

The reason behind making Speech-Text Conversions of paragraphs is it is one of the most popular use cases and other is that it built knowledge about which module or package is available especially when get chance to work with big projects. One more reason is that, it helps me to learn and communicate in various languages with my friends who are staying abroad. It provides ease of communication especially travel as tourist to other countries. Since they are pre-built libraries provides most accurate and reliable results with very few errors. Its secure pathways for information transmission. In short, it provides flexibility to work in and out of the organization. Text to speech helps deaf, lazy or people who prefer to listen more rather than reading. Translated text helps to learn many languages digitally without paying and also helps to communicate with foreigners. By building such kind of apps, I have started to learn German, Dutch and Swedish. It also helps to understand their culture and thoughts. It is proven fact that person who can speak more than one language have improved memory, enhanced concentration, problem solving and critical thinking skills. They can even effectively participate in multi-cultural world.

Technical Aspect

Speech Recognition can produce documents in less than half the time it takes to type. It is the process of converting spoken words to text. It converts audio signal into a spectrogram which is 3D graph. Its analysis each sound using algorithm to find the most probable word fit in that language and transcribing those sounds into text.
googletrans uses Google Translate Ajax API to make calls to such methods as detect and translate. It is fast and reliable. It has auto language detection feature. It can do bulk translations. It is customizable service URL. It generates a ticket by reverse engineering.
gTTS is Google Text-to-Speech module, CLI tool to interface with gTTS API. It is customizable speech-specific sentence tokenizer that allows for unlimited lengths of text to be read. It has feature of automatic retrieval of supported languages.
pydub manipulate audio. It helps to perform all audio processing tasks such as loading, saving, appending, mixing, changing audio levels generating tones etc.

Installation

Using intel core i5 9th generation with NVIDIA GFORECE GTX1650.
Windows 10 Environment Used.
Already Installed Anaconda Navigator for Python 3.x
The Code is written in Python 3.8.
If you don't have Python installed then please install Anaconda Navigator from its official site.
If you are using a lower version of Python you can upgrade using the pip package, ensuring you have the latest version of pip, *python* -m pip install --*upgrade pip and press Enter*.

Run/How to Use/Steps

Keep your internet connection on while running or accessing files and throughout too.
Follow this when you want to perform from scratch.

Open Anaconda Prompt, Perform the following steps:

Creating Virtual Environment named “smartvoice”. You can give any name of your choice.

conda create -n smartvoice python=3.6

y

conda activate smartvoice
pip install SpeechRecognition
pip install pydub
pip install googletrans
pip install gTTS

cd <PATH>

You can also create requirement.txt file as, pip freeze > requirements.txt

run files.

Creating Virtual Environment is necessary so that you do not have to install packages every-time you run the code. Once all required packages are installed in virtual environment then you only need to access/open the virtual environment and run the final file.

Follow this when you want to just perform on local machine.
Download ZIP File.
Right-Click on ZIP file in download section and select Extract file option, which will unzip file.
Move unzip folder to desired folder/location be it D drive or desktop etc.
Open Anaconda Prompt, write cd <PATH> and press Enter.
eg: cd C:\Users\Monica\Desktop\Projects\Python Projects 1\SpeechRecognition\largefiles_Conversion
Now, open virtual environment that you have created ie
conda activate smartvoice
In Anconda Prompt, pip install -r requirements.txt to install all packages.
In Anaconda Prompt, write python <filename>.py and press Enter. That is,
In Anaconda Prompt, write python 1)speech_to_text.py and press Enter.
This takes vacation_audio.wav file as input and creates recognized.txt file as output text.
In Anaconda Prompt, write python 2)text_to_translate.py and press Enter.
This takes recognized.txt file as input and creates recognized_2.txt file as output text.
In Anaconda Prompt, write python 3)translated_text_to_speech.py and press Enter.
This takes recognized_2.txt file as input and creates audio_2.mp3 file as output audio.
You can see audio-chunks folder getting created in working directory.
You can also run all codes from Command Prompt instead of Anaconda Prompt after setting Environmental Variable Path Settings.

Note: I have created smartvoice virtual environment and used for more than one project and therefore you might see more than one unused library in requirements.txt especially for this

project so do not worry because I am using them in another project under similar virtual environment. Whenever you get No Module <name of package> Error then see its PyPI Documentation and Install it using pip install <package-name> written there. In some cases, you need to install its .whl file which I will inform you if its necessary.
Please be careful with spellings or numbers while typing filename and easier is just copy filename and then run it to avoid any silly errors.
Note: cd <PATH>
[Go to Folder where file is. Select the path from top and right-click and select copy option and paste it next to cd one space <path> and press enter, then you can access all files of that folder] [cd means change directory]

Directory Tree/Structure of Project

Folder: SpeechRecognition>largefiles_Conversion
1)speech_to_text.py
2)text_to_translate.py
3)translated_text_to_speech.py

To Do/Future Scope

Can add augmented reality feature however yet to dive in.

Technologies Used/System Requirement/Tech Stack

Credits

Medium.com