ReadMe

Company ---- HardHat Enterprises

Project ----- Breaking Audio Captcha Using GCP Speech-To-Text API

Trimester 3, 2022

Table of Contents

1. Project Information

1.1. Company Acting Director

Anh Dinh
Senior Lecturer, Cyber Security
anh.dinh@deakin.edu.au

1.2. Project Team

Project Name: **Breaking Audio CAPTCHA Using GCP Speech-To-Text API**

Company: **Hardhat Enterprises**

Team:

Name	Role
Ridwan Ganiyu	Audio CAPTCHA R&D

2. Project Overview

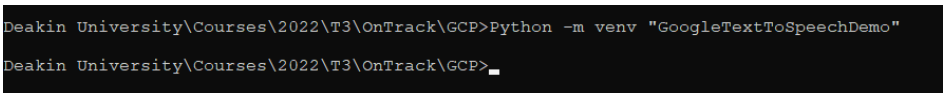
The goal of the umbrella project - Breaking CAPTCHA project, is to develop machine learning, AI, and computer vision models to autonomously solve CAPTCHA problems. The final product of the project will be to provide end users with an interface to solve these CAPTCHA problems. The product will also be a pipeline that Developers can use to collaborate on ML experiments. Both these user interfaces (users and ML developers) will be the focus of future iterations.

The goal of Breaking Audio CAPTCHA project is to develop a AI/ML solution that breaks audio captcha, hence this ReadMe is a result of the aim to produce a solution that does that. This project uses AI/ML algorithm provided by Google Cloud Services to break reCAPTCHA and BotDetect captchas.

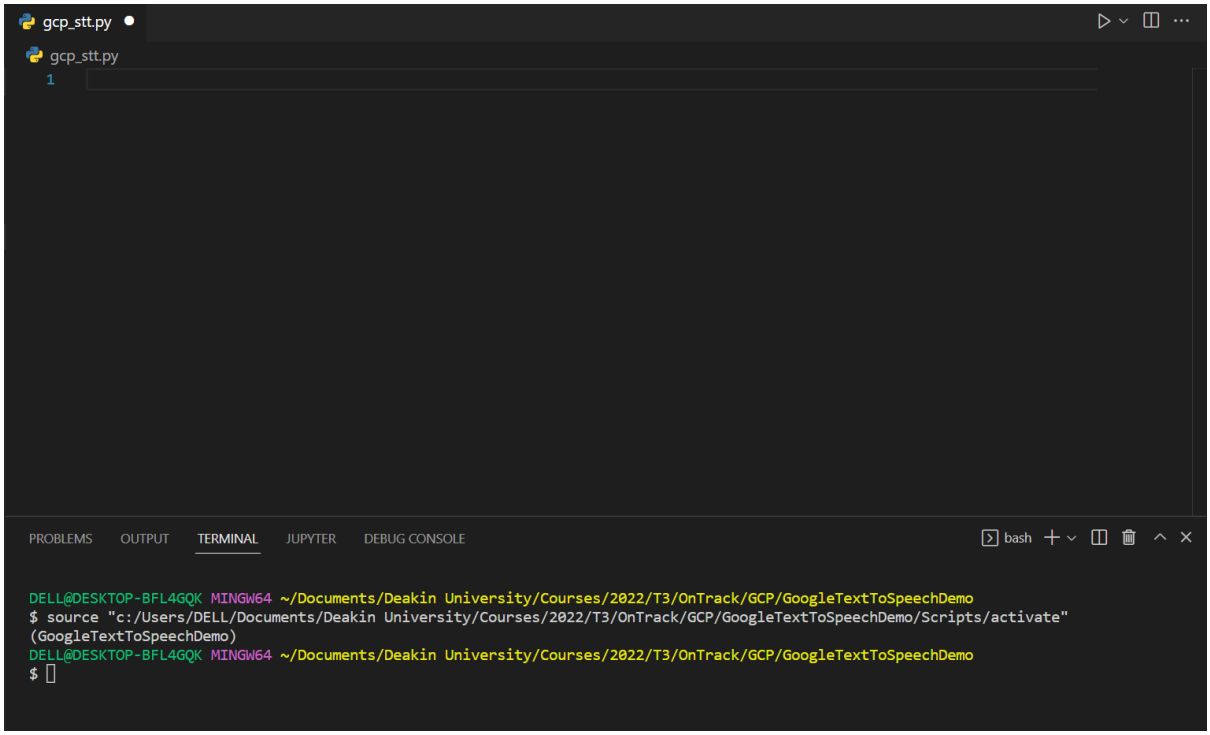
The deliverable of this project is code to demonstrate that GCP speech-to-text can break both captchas using the applicable ML model.

3. User Manual

In the use of the API, I used the 'Visual Studio Code' application as the environment, installed the latest python application my device, and then created a python environment from 'cmd':



Use Visual Studio Code to open the created environment, create python file, set the python interpreter, and then run the blank python file.



This should be recognized in the VS code terminal as shown above.

Install the client library into the created environment via the terminal, to be able to use the API:

pip install –upgrade google-cloud-speech

- Using GCP Speech-To-Text API:

This part is used to setup an account on GCP, the environment and API service in Google Cloud to receive and send requests:

- Create a GCP project in your service account and enable billing for the project.
- Enable Speech-to-Text API on the project.
- Create a service account credential key and set Authentication Environment Variable.
- Create a Cloud Storage bucket to store Audio data
- Enable ‘Data Logging’ to record audio data sent to STT for improving the STT model. (optional)

Given that a service account and key has been created, we can go ahead to using the API. The ‘gcp_recaptcha.py’ file contains python code to use on reCAPTCHA, while the ‘gcp_botdetect’ contains python code to use on BotDetect captchas.

4. Completed Deliverables

Completed Deliverables for this project include the following:

- Review last trimester report:
  reviewed the last trimester report on configuration

- Research GCP speech-to-text
- Configure and use API on sample audio for speech recognition
- Configure and use GCP API on reCAPTCHA
- Configure and use GCP API on BotDetect captcha

Below are links to Trello and Github repo:

- https://trello.com/b/vEd7OIDZ/bc-t3
- https://github.com/Hardhat-Enterprises/breaking-captcha/tree/main/Teams_T3_2022/Audio_CAPTCHA/Ridwan_Ganiyu_Contribution

5. Roadmap

This novel project creates an alternative to the already developed solution based on transformer-based ML model.

Further deliverables of this project will be:

- Further testing on reCAPTCHA and BotDetect captchas to further tune the API use
- its integration into a toolbar solution and the web based breaking captcha development

6. Login Credentials

The credentials required for this is the service account key created for the project.

7. References

Below are references used in the preparations, research, and development:

- Google Cloud, Make an audio transcription request, Cloud Speech-To-Text, December 2, 2022
- Google Cloud, Recognize speech by using enhanced models, Cloud Speech-To-Text, December 2, 2022
- Google Cloud, Client libraries, Cloud Speech-To-Text, November 30, 2022
- Google Cloud, Select a transcription model, Cloud Speech-To-Text, December 5, 2022
- Boosting Speech-to-Text API accuracy, Google Cloud Tech, AI and Machine Learning with Google Cloud, September 14, 2022