Journal Report 2
9/8/19-9/14/19
Arvind Ravipati
Computer Systems Research Lab
Period 1, White

Daily Log

Monday September 9

Looked into Different Speech to Text APIs from https://nordicapis.com/5-best-speech-to-text-apis/. Looked at which of them Costed Money and which had Free Trials and were Implemented in Python. Found Speech Recognition and Found that it was Free and Implemented in Python

Tuesday September 10

Looked at the Tutorial from https://realpython.com/python-speech-recognition/ to Start Coding. Installed Necessary APIs and Tried Implementing with Sample .wav Files. Script Kept Returning an Error, so I Looked through Different Stack Overflow and Github Threads. Found that Method Needed More Parameters to Work Optimally. Script Started to Return Empty Arrays so Had to Research Better Options or Potential Fixes.

Thursday September 12

Found that I had to Reduce all Background Noises from the Sound Files for Speech Recognition to Work Best. Found Shorter .wav Files with Reduced Background Noises and it Worked. However, Different .wav Files did not Work, so I Assumed it had to do With the Type of Voice and the Amount of Background Noise. Found out how to Download .wav Files for Each TED Talk.

Timeline

Date	Goal	Met
Aug 30	Finalize Idea and Think of Good Starting Point for Project	Yes, Submitted Journal 0
Sep 6	Finish Formatting Sizeable Dataset	Yes, Dataset has 2461 Entries
Sep 13	Find Good Audio to Text API or Github Repository and Implement	Yes, Found Speech Recognition and Tested with Sample .wav File
Sep 20	Be Able to Input a Whole Ted Talk and List the Most Probable Transcript	
Sep 27	Find Best Extension of Speech Recognition and Compare Transcripts	

Reflection

Many of the speech to text APIs required money, so it was good to find one that was free and implemented using Python. Initially, my script was returning errors or empty arrays so I thought I would have to look into different options. After trying a different sound file, it worked and returned an accurate transcript. In addition, it was a bonus to see the confidence as it could be later used for results and comparing different transcripts of the same audio file. I also thought that something complicated had to be done to download the audio from the TED talks, but found that podcasts could be downloaded. Below is a snapshot of the result from the working sample run.

```
Arvinds-MacBook-Pro:Sys Lab arvindravipati1$ python3 speech_to_text.py
{'alternative': [{'transcript': "you're so funny", 'confidence': 0.9876
2912}], 'final': True}
```