CS410 Course Project – Progress Report

1 Team Information

Team name: Torrey Pines
Team member (NetID): Wusi Chen (wusic2)
Team captain: Wusi Chen

2 Finished Tasks

- Parse subtitles files (WebVTT files) of CS410 course videos, and store the parsed data in a CSV file and a compatible file to Metapy.
- Build a method in Python to return top-relevant transcript segments based on the queries.
- Build a search engine website for video segment search.

3 Remining Tasks

- Verify the parsed data of the video segments from subtitle files:
  - Check if the transcript of the video segments is complete and correct.
  - Check if the start timestamp and end timestamp of video segments are correct.
- Verify the search result (a list of relevant video segments):
  - Check if the links to video segments are correct (play the lecture video at the correct start timestamps of the video segments).
  - Check if the returned video segments from the program are related to the queries.

4 Issues/Challenges

- Flask was initially used as the web framework for the search engine application, but the Python program gets stuck and runs forever when call the it calls "ranker.score" method. To solve this issue, another Python-based web framework, Bottle, was used as the web framework for the application.