INTELLIGENT LEARNING PLATFORM

CS 410 : Project Proposal

Prepared for: CS 410

Prepared by: Aparna Anand

24 October 2021

Team Members

This will be a single member team. Aparna Anand (aparnaa2@illinois.edu) will be responsible for building and delivering the project and the milestones on time.

Topic

The topic chosen is from the Intelligent Learning Platform. The goal is to build a search engine that allows for video segment search.

Project Outline

All have faced a situation where we would like to have the ability to search through a video to just a particular segment of interest or search through our video archive to find a few specific moments. As mentioned in the Topic section, this project intends to create a search engine that will allow video segment search.

For this to be possible, the videos need to be enriched with metadata. Metadata is not just the date, location and rich description but also a transcript of the video. If we could annotate any object, Text, or signs that appear in the videos, we could search for specific moments when they occur.

In the project, the dataset used will be 1-2 lecture videos of class CS 410, 1-2 random videos from my video archive. I want to use the Google video intelligence API to annotate the videos from my archive using Label detection. For the lecture videos, I plan to use the API to detect Text, and I also want to store the transcription to help with search of segments from the video.

The project proposed is relevant to what we are learning is because of questions such as

- Which video segment is most relevant to the search text?
- Have all the relevant video segments been retrieved?
- Are they ranked correctly, or is there a ranking required?

We have gone in-depth about the questions pointed out above for text-based search engines. The same applies by annotating the videos and allowing users to use Text to search relevant annotated video segments.

The language I will use is Python. For using the API and getting the annotation results, I will need a google cloud account. As for displaying the results, I am still not very sure what to use. One of the links I have provided in the reference talks about using streamline.io, which is one of my considerations or I may build a simple webpage to display the results.

Time

Most of the content discussed above in the project outline section is new to me, and I will have to explore how the labelling works, storing the labels and building a simple search to pick the relevant video segments. Finally, once all this is in place, a proper visualisation to display the results. Being a single person team, this will take me 20+ hours.

An assumed split would be

Exploring the API and how video annotation works:

1. Initial system setup ~2 hours
2. Label detection ~2-3 hours
3. Text detection ~2-3 hours
4. Transcript usage + labels from Text detection combined to enrich search ~ 2-3 hours
5. Retrieving and storing the labels and time segments ~2 hours
6. Simple search to use text and retrieve results(includes ranking and bringing relevant information) ~5-8 hours
7. visualisation of the results ~5 hours
8. Final project demo video and report ~2- 3 hours

References

https://mlbhanuyerra.github.io/2019-12-09-Video-Search-Engine-Salsa/

https://towardsdatascience.com/prototyping-my-video-search-engine-d6fb03c9bcd1

https://www.erikaagostinelli.com/post/getting-started-with-google-video-intelligence-api-using-python

https://cloud.google.com/video-intelligence/docs/analyze-labels