

Project Proposal jkchang2

1. What are the names and NetIDs of all your team members? Who is the captain? The captain will have more administrative duties than team members.
 - a. Jonathan Chang, jkchang2, captain
2. What topic have you chosen? Why is it a problem? How does it relate to the theme and to the class?
 - a. Intelligent learning platform. It is becoming increasingly common to learn online as opposed to a traditional classroom experience. This can be for a variety of reasons, including flexibility, cost, and individual preference. Specifically, I'm interested in text within the video lecture captions. There is a lot of information and keywords captured by lecturers, which are not sufficient to understand from the slides alone. To improve the student's learning experience, it would be great to highlight uncommon words used as keywords, tag them to videos, and allow for easier reviewing of material. Otherwise, most times students have to click through videos and search for the content they are looking for manually. This will turn the raw data into actionable insights for learners to optimize time spent learning instead of wondering where the information is located.
3. Briefly describe any datasets, algorithms or techniques you plan to use
 - a. The raw data is captured by the captioning system Coursera provides. It would be possible to scan that text, apply something as simple as TF-IDF, and highlight keywords, as well as store them to be viewed alongside each video (TBD on optimal UI experience).
4. How will you demonstrate that your approach will work as expected? Which programming language do you plan to use?
 - a. We can do a demo by taking a sample coursera lecture page, scraping the text, and showing the overlay via HTML/CSS/Javascript. Ideally the keywords can be organized and shown in a summary format.
5. Please justify that the workload of your topic is at least $20 \cdot N$ hours, N being the total number of students in your team. You may list the main tasks to be completed, and the estimated time cost for each task.
 - a. Documenting the ideal user experience (is it highlighting keywords, can we call them out somehow in the navigation bar alongside videos?). Provide walkthrough of this and some visuals.
 - i. 3 hours
 - b. Testing some different methods to implement given short time to execute (TF-IDF, or otherwise)
 - i. 5 hours
 - c. Complete the actual source code with testing
 - i. 7 hours
 - d. Prepare the demo to be shown (ideally directly from the actual page itself, or as a standalone work for full functionality). Need to test out limitations and feasibility
 - i. 5 hours