

CS410 Course Project

Group Members:

Rijul Tikku (rtikku2) - Captain

Description:

For the course project, I am planning on creating a chrome extension intended to be used with any text-based webpage that would contain questions. This includes but is not limited to forums like Reddit, Stack Overflow, and Quora. Once the extension is activated, the webpage will be scraped to look for all relevant questions posed on the page and answer them one by one in the extension dropdown using the context of the page as well. From here the user can navigate to other web pages that may contain the information that they are looking for.

Topic and Problem:

The topic that I chose for this project was Intelligent browsing. This is an extension that I am adding to Google Chrome to extend its browsing capabilities. The problem this solves is when a user finds a webpage where a question is posed that aligns with what the user is looking for, but the answer might not be quite right. There might be a question asked on a forum that is unanswered but provides a lot more context to the question than just simply the query. Here, the user will be benefited by being able to find similar resources where the question is actually answered. In other cases, a user might find just the question they are looking for, but the answer isn't what they hoped for; it could be that the answer isn't thorough enough, hard to understand, or simply incorrect. This relates to the theme of the class as it will use the information already present on the website to try to help the user browse the internet in a more efficient manner.

Datasets/Algorithms/Techniques:

I will need to use some algorithms in order to determine when questions are asked and what information on that page could help bolster the query to find better answers to the questions on the page. Some methods that initially stick out to me that would be beneficial for this project include BM25 to rank the possible results that the extension will return and content-based filtering to find appropriate results.

Approach:

In order to demonstrate that my project works as intended, I can go on various forums that have questions that either don't have answers or incomplete answers and use my extension to find results that answer the questions asked more thoroughly.

Programming Languages:

In order to make a Chrome Extension, you will need to use HTML, CSS, and Javascript and those are the languages that I will be using.

Workload:

Since I am working on this project alone, I believe this project will take over 20 hours to accomplish through the semester. I will first need to figure out how to get a chrome extension up and running. I don't have much experience with front-end languages so I expect there to be a

learning curve when working with them. If I had to estimate I would expect to spend 8-10 hours on that aspect. After that, I will need to get the text scraping and question recognition working. I can estimate spending 3-5 hours to get this working. Finally, I will need to use that scraped data to find appropriate results. I estimate this to take the longest amount of time at around 10-15 hours.