CS 410 Fall 2021

Authors: Peter Hartman, Steven Hernandez, SangJin Hyun

Project Proposal

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.**

* Peter Hartman (NetID: pehartma)
* Steven Hernandez (NetID: stevenh5)
* SangJin Hyun (NetID: sangjin2)

1. **What topic have you chosen? Why is it a problem? How does it relate to the theme and to the class?**

Our group project follows the *Intelligent Browsing* topic, by aiming to provide a browser extension for Chrome to highlight keywords in webpages and transform them into hyperlinks which the user can click to search more information on the keyword. This extension would use text retrieval techniques to rank candidate keywords and exclude stop words, thus providing a user like a researcher with a convenient method to improve research efficiency. It would also provide controls to limit the number of keywords as well as search for a keyword.

1. **Briefly describe any datasets, algorithms or techniques you plan to use.**

For this project, we will be using text retrieval to match user query (for keyword search) with the VSM algorithm with stop word elimination, page indexing, and text mining using Wink-NLP library for JavaScript.

1. **How will you demonstrate that your approach will work as expected?**

Our group will be taking on the role of user “experts” to judge the quality of highlighted keywords, as well as confirm that tuning controls respond as expected and that the keywords are correctly transformed into hyperlinks and can be operated with well.

1. **Which programming language do you plan to use?**

JavaScript, (Python)

1. **Please justify that the workload of your topic is at least 20\*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.**

|  |  |  |  |
| --- | --- | --- | --- |
| Member | Task | Estimated Hours | Total (60 hours) |
| Peter Hartman | Program structure, beginning script, UI, display and execute module, module integration | 24 | 24 |
| Steven Hernandez |  |  |  |
| SangJin Hyun | Work on program design, parsing, computational module | 20 | 20 |