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.

pkim62, bsutter2, bsutter2 is the captain!

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

We chose the Intelligent Browsing topic and we have chosen to create a browser extension to allow you to search based on concept rather than exact words. The current system requires exact spelling and words for searching using “ctrl-f” but we would like to allow synonyms to be recognized together as one concept for searching. This uses text mining and analysis to provide easier searching.

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

We plan on using a common retrieval function like BM25 to index a document to improve the searches over a particular page. We may need a dataset to provide certain information for lexical analysis information.

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

We will show that our approach works as expected by illustrating a better understanding of core concepts in the course and providing a functional product that is an improvement of current extensions in chrome. The approach should increase precision of the ctrl-f function in chrome and reduce the number of useless words.

5. Which programming language do you plan to use?

Python

6. 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.

With only two members of our group, this project will likely take well over the given 40 hours we should spend. We must learn how to create chrome extensions, create a program that uses text mining and analytical algorithms along with pooling algorithms to group common terms.