Project Proposal

Topic: Chrome “Find in webpage” with auto-correction, stemming, and synonym detection

Team members: Xiaohe Cheng (xiaohec3@illinois.edu) - Project coordinator

Problem statement: Search within a webpage (“Find”, Ctrl+F) is a very common use case in web browsing. Chrome, one of the most popular web browsers, provides built-in search which finds all occurrences that precisely match user input. However, user input is not always precise. Users sometimes make typing mistakes. Sometimes they want to go back to a paragraph they have read but cannot recall the exact content. Sometimes they only have a rough direction of what is desirable. To cater for this need, this project proposes to build a Chrome extension that enhances Chrome “Find in webpage”, by adding the following features:

- Auto-correction, e.g. user types “Probablistic”, also search for “Probabilistic”
- Stemming, e.g. user types “Creating”, also search for “Created”
- Synonym detection, e.g. user types “Novel”, also search for “Innovative”

This project is highly related to CS410, because tasks like stemming are classic NLP problems.

Methodology: The plan is to use JavaScript to write a Chrome extension. Considering the limited time and resources, it is desirable to leverage existing libraries as much as possible. For the three tasks, the npm autocorrect, stemmer, and synonyms package can be good start points. If time allows, I can customize or extend the libraries.

Deliverables: A working Chrome extension, with examples that demonstrate the features.

Workload analysis:

Task	Estimated time
Set up a Chrome extension that can take user input, read HTML, and control (e.g. highlight the search results)	8 hours
Implement and test basic auto-correction	3 hours
Implement and test basic stemming	3 hours
Implement and test basic synonym replacement	5 hours
Test and further improvement	5 hours