CS410 Fall 2021

Authors: Peter Hartman, Steven Hernandez, SangJin Hyun

Progress Report

1) Which tasks have been completed?

The core Chrome extension framework has been written, including the popup functionality with the control slider to adjust the number of terms to highlight. The code for replacing text within the HTML documents with search hyperlinks is finished and only requires minor testing. A simplified version of the document processing model is complete and is capable of tokenizing the input HTML body, converting it to a bag of words, and selecting the top terms by frequency.

2) Which tasks are pending?

The model needs to be enhanced with preprocessing to remove stopwords. Querying the model for an alternative ranking of tokens is not yet supported and will require an additional element in the extension popup. User testing of the extension once feature complete is required.

3) Are you facing any challenges?

There have not been any major challenges at this time.