Captain: Joseph Jang (jyj3)

CS 410 Project Proposal

1. What topic have you chosen? Why is it a problem? How does it relate to the theme and to the class?
    a. I have chosen to build a google extension that performs sentiment analysis on user selected text, and highlighting or blocking the content. Considering the massive amount of text online, this tool can help users to quickly filter contents and for use in emotional shielding. This project falls under Intelligent Browsing, giving an intelligent insight into understanding browser content, and helps users to better navigate browsing by quickly filtering out information they may not want to see.
2. Briefly describe any datasets, algorithms or techniques you plan to use
    a. I plan to explore different directions in text sentiment analysis, such as libraries that employ lexicon-based methods that utilize predefined weighted positive/negative words, or machine learning based methods like linear regression/SVMs training on large labeled sets, or pre-trained language models like BERT and chatGPT for sentiment analysis. I will mainly focus on different node.js NLP modules that build on these methods.
3. How will you demonstrate that your approach will work as expected?
    a. I will use different diverse web pages with various sentiments, such as movie reviews with ratings, social media comments. By testing the extension on these pages, I can verify the correct sentiment analysis.
4. Which programming language do you plan to use?
    a. I will use Javascript and HTML/CSS as these are the main languages for web based projects and particularly google extensions. These will allow me to interact with the DOM of the webpages to extract the text I will be performing sentiment analysis on.
5. 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.
    a. These are the minimum estimates

b. Research & Planning (3 hours): Understand sentiment analysis, how to build a google extension, and website DOM interactions
c. Development of Extension (7 hours): Development + sentiment analysis.
d. UI/UX Implementation (3 hours): Interaction with DOM / displaying results
e. Testing & Debugging (5 hours): RIP
f. Documentation & Tutorial (2 hour)