CS410 Project Proposal
Christopher Dimitri Sastropranoto 678097021

Project Description
The proposed project is to implement a google chrome extension that can quickly summarize a Medium articles.

Who will benefit from such a tool?
Almost anyone can benefit from such a tool. For example, I am an avid user of Medium but feel that some articles are unnecessarily lengthy and would love to save some time just getting the gist of an article by having a quick summary of what it is about. If I'm interested in the summary then I could go ahead and read the full article.

Existing Tools
Doing a quick google search resulted in several chrome extensions and websites that do something similar. The biggest drawback from most of them was that they required the user to either enter in a link to an article, or copy and paste the article's contents. The goal is for my tool is to allow the user to skip those steps and simply click a button to summarize the article's contents whilst on the browsing.

Programming Language and other Other resources
There are currently several online services that I can use to help implement this project. My plan is to use the **python NLTK library** to help with summarizing text and packaging that inside a Chrome Extension UI. In addition to this, I plan to do my own online research by utilizing YouTube, Medium and other websites to augment what I've learnt in CS410.

Techniques and Algorithms
From my research so far, summarizing text requires some preprocessing steps such as lemmatization and POS tagging. Once this is done, the next step would be to model topic distributions in the document. In addition to this, there are several summarization techniques that need to be explored. These are Extraction-based summarization, Extractive summary and Abstract based summarization.

How will you demonstrate the usefulness of your tool?
If successful, I plan to demonstrate the usefulness of my tool by creating a video and also releasing it in the Google Extension store.

Project timeline

Project start date: 10/26/20

Milestone	Target Date
Complete research into summarization techniques	10/28
Preprocess Data	11/4
Train text summarizer. This step will be done in isolation using JupyterNotebook.	11/8
Build chrome extension and hook up to text summarizer. This step requires implementing a way for the extension to scrape the article's contents from the website.	11/13
Testing and Bugfixes if needed	11/16

Work Justification

I feel that the scope of the project sufficiently fulfills the 20 hour requirement as it involves multiple components. This includes doing further research into the summarization techniques, preprocessing the data, training it using NLTK and finally implementing a Chrome Extension UI for it.