Targeted Information Learner

Track: Development

Team members:
- Jiadong Gui, jgui3@illinois.edu
- Ray Ko, wko21@illlnois.edu
- Ruiying He, rhe10@illinois.edu
- David Lin, davidl16@illinois.edu

Coordinator: David Lin

Functions and Users

Either chrome extension of web app that takes in the contents of a web page summarizes information on the web page, looks at hyperlinked sites on the page, and records any images/diagrams on the page. Based on this information, the app will propose new ideas to explore, areas to be more in depth on, or users can create their own query. For example, on a page about bfs with diagrams and links to a dfs article, the app may ask whether you want to learn more about how bfs is implemented, learn about dfs, explain clearly what is happening in a specific diagram, or propose your own question to ask. The users of this tool will mostly be high school and college students, but it can be used by anyone who wants to learn more about any topic. Our app can be used to digest large amounts of information like if a student is reading a research paper.

Significance:

This tool is needed because existing technology only summarizes such information but does not enable users to actively engage with information. Our tool will help users learn and digest new information by actively engaging with the content and open up new paths by exploring new related information. A potential pain point we address is being overwhelmed with information when learning new things. This tool will allow users to summarize information and then target key points they feel need a deeper explanation or need more clarity on. This allows users to focus on the important or most interesting bits of information while quickly learning less relevant things. It also helps speed up certain processes as right now in order to analyze images, learners generally need to switch between screenshotting the article and an LLM. Our tool erases this dependency and has all images stored and prepared for analysis.

Approach

In our approach, we would use a front end interface built with HTML/CSS and Javascript that the user can interact with. Users would be able to send a url to the website of their choice, and our backend service, built with Flask, will feed information from the site into our LLM. Any articles hyperlinked on the site will also be fed into our LLM, and embedded images will be extracted and encoded to base 64 to prepare for analysis with gpt-4-vision-preview, which allows for image analysis. Our LLM will be pre-trained to ask follow up questions related to the article, hyperlinked sites, and embedded images.

Evaluation

We will evaluate our app by feeding in articles and seeing how relevant the follow up questions are to the contents of the article, related hyperlinked information, and embedded images. We will ask other people to use our app and answer surveys detailing our apps, quality of life, intuitiveness, whether it sped up or slowed down the learning process, increased learning retainment and understanding, and overall usefulness.

Timeline

- **Week 1** - Setup an interactive frontend interface for submitting articles/link and allowing user to input their own queries
- **Week 2** - Take articles and web pages and scrape the data for the text body, hyperlinked sites, and any embedded images.
- **Week 3** - Setup LLM and integrate it with the backend. Ensure data such as text and images can be fed into the LLM and produce relevant results
- **Week 4** - Customize the LLM so it can follow up with the user about any areas to explain in more detail, related topics, or explain any images.
- **Week 5** - Test, evaluate, debug, and modify with any final changes to improve app

Task Division

- Jiadong Gui
  - Work on an interactive and intuitive front end
- Ruiying He
  - Scrape articles for information, text, and images
- Ray Ko
  - Work on backend and integrate with frontend
- David Lin
  - Setup LLM api for images and text