UR2PhD Proposal

Lateral Reading

Anna Li and Diana Brebeanu

Cheriton School of Computer Science
University of Waterloo
Canada
December 7, 2023

Proposal

Research context and problem statement

Online misinformation is a pressing concern, and addressing its proliferation requires a multifaceted approach. One method to combat this issue is the system-focused approach, which involves the censorship of malicious news sources. However, this approach raises significant concerns regarding the preservation of freedom of speech. A more promising alternative is the individual approach, which centers on the way individual readers consume information. One key strategy within the individual approach is lateral reading, which entails engaging with an article, asking critical questions about the information presented, seeking answers online, and summarizing these answers. In short, Lateral reading is using multiple sources of the same subject to validate and verify the topic, rather than reading in only one source. Therefore, lateral reading serves as a valuable tool to assist readers in assessing the credibility of the articles they encounter. [1] Lateral reading was devised after researching the strategies of professional fact checkers compared with historians and students. It was found that while the students and historians remained on the original website for longer to detect its reliability, the fact checkers quickly opened multiple tabs to verify the information, coming to credibility conclusions in a fraction of the time. [2] Therefore, this act of quick verification by cross-referencing other sources has been deemed lateral reading. Our current research initiative is dedicated to developing a conference where researchers are encouraged to devise various implementations of lateral reading tools. To test and evaluate the effectiveness of these new tools, we need a test set of articles that contain complex language and potentially contentious content, ensuring that the tools can effectively analyze such articles. Our central challenge, therefore, lies in defining an appropriate evaluation dataset to rigorously test the machine learning models underpinning these tools and provide a robust analysis of the articles in question.

Proposed Solution

To address the complex issue of evaluating the effectiveness of lateral reading tools, our team has deliberated and established a strategy. After thinking the criteria and discussing with our mentor, we've delineated essential components that should be present for our test articles: they should necessitate lateral reading, encompass a diverse range of subjects from politics, human interest, sports, business, opinion, science, health, arts, and food, and be sourced from the period starting from 2021 until the end of February 2022. Each team member has been tasked with identifying 40 fitting stories, diligently recording key details such as the story name, associated wiki page, event date, and a description of the narrative. After individual selections, we will consolidate our findings and, through careful analysis, use these stories to refine our collection down to 50 quintessential stories. These will then serve as the basis for which to find articles based on these stories to create the test set to be used against the lateral reading tools, providing a comprehensive evaluation of their capability and effectiveness.

In our effort to extract relevant articles based on these stories from news media sources, the following steps will be employed. First a list of global news media sources

will be compiled. This data will be systematically arranged in a spreadsheet, with columns dedicated to the source_name, domain_name, and bias_rating. The bias ratings of these media sources will be used to compare with the articles found based on the stories. Subsequent to this, we will implement queries using a retrieval algorithm on the Compute Canada general-purpose cluster to retrieve the appropriate articles using queries based on the relevant stories found. The final step involves the organization of the retrieved data. A Google spreadsheet will be maintained, listing the doc_ids, queries, and story_id. For comprehensive representation, it is advised to have a minimum of five articles associated with each story with a balanced proportion of bias ratings.

Our research is based on the Dake Zhang and Ronak Pradeep's previous research Readprobe [1].

Evaluation and Implementation Plan

Evaluation Plan

Since the solution is made of several different parts, we have corresponding different evaluations.

1. Researching important news stories from the 2021 up to Feb 2022:
   We each categorize the comments of different articles individually, and then we calculate the Cohen's kappa's constant to check our consistently and the quality of our categories.

2. Compiling the database of global news sources and their biases:
   As this is done through a process of collecting data, then proper evaluation consists of looking through the dataset and checking that the code was able to parse and group all the data in the correct categories with no elements missing.

3. Running the queries on the crawled news articles:
   The effectiveness of the query can be evaluated by reviewing the content of the top 30 articles returned to see how relevant the article is to the story we are searching for. We also review the BM25 score returned by the model for each article which states how relevant the article is to the query. If the score is too low or not enough articles are relevant, we change the query by changing the key words and which words are emphasized.

4. Database of articles with balanced media biased:
   We can evaluate the suitability of the dataset to be tested on lateral reading tools by calculating the proportion of articles whose stories lie in each story category (politics, human interest, sports, business, opinion, science, health, arts, and food) as well as the proportion of the media biases (leaning left, leaning right, and centre). If these proportions are equal between all of them, then this would be a suitable dataset. Our end goal is to have 6 articles per news story with 2 that are left-leaning, 2 that are center-leaning and 2 that are right-leaning.

Timeline

We are completing this research over the course of a 4-month term. The remaining time left is for the month of November which gives the following timeline:

Oct 30 - Nov 6:
Compile a database of global news sources and their biases.

Nov 7 - Nov 13:
Learn and develop queries to be used with the retrieval algorithm to return relevant articles from the crawled article database.

Nov 14 - Nov 20:
Compare relevant articles with the bias of their sources to select a number of relevant articles with an equal proportion of biases.

Nov 21 - Nov 27:
Review and refine database of articles with supervisors with any extra work being allotted during this and the next week.

Nov 28 - Dec 4:
Finish any remaining work as well as finish the research proposal and record the presentation.

References

[1] Ronak Pradeep Dake Zhang. Readprobe: A demo of retrieval-enhanced large language models to support lateral reading. June 2023.

[2] McGrew Wineburg, S. Lateral reading and the nature of expertise: Reading less and learning more when evaluating digital information. *Teachers College Record*, 2019.

Revisions

The revisions made include explaining the key terms some more in the Context section and Proposed solution sections. We also modified the evaluation plan to align more closely with the changes in our research. We also changed the formatting of the proposal and added a title page to make it look more formal.