nut-kslab at TREC 2023 CrisisFACTS track

Phichamon Theamtun [1], Takashi YUKAWA [2]

[1]
Nagaoka University of Technology, Niigata, Japan
s225076@stn.nagaokaut.ac.jp

[2]
Nagaoka University of Technology, Niigata, Japan
yukawa@vos.nagaokaut.ac.jp

1. Introduction

This notebook is the summary of our approach for the TREC CrisisFACTS 2023. Our approach will be presented in Section 2, and our run and results will be discussed in Section 3. With run discussion and problem we faced led to our future work in Sections 4 and 5.

2. Approach

To generate summarization lists for each day of the 18 events, we began with a ranking function used in information retrieval and text search, BM25. Our decision to employ this approach was based on its established relevance to the field, despite our first-time application in this context. However, the results turned out to be inadequate.

3. Run and Results
3.1 Run setup

A retriever with BM25 as the weighting model is set up. We use dataset queries to find relevant documents and save the scores. Then assesses document importance by counting how many times they are deemed important.
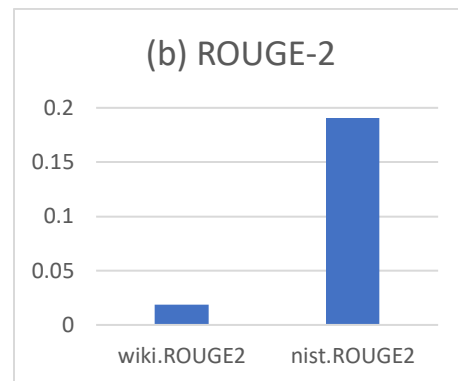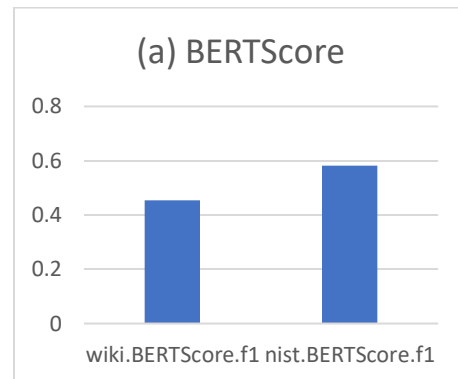
3.2 Results

Figure 1: Evaluation results (BERTScore) of our submission. Figure 2: Evaluation results (ROUGE-2) of our submission

Evaluation of our submission using denote summaries extracted from NIST-based assessments and Wikipedia articles. Figure 1 and Figure 2 show BM25 run is more semantically similar to the NIST reference than the Wikipedia reference.

4. Consideration

Relying solely on the BM25 method was an intentional choice, given its wide acceptance in the information retrieval field. However, as the research progresses, there is a contemplation of advancing to more refined approaches, such as employing GPT-4 with more suitable components for future runs.

5. Summary and Future Work

While considering the count of occurrences and the sum of importance scores, the latter option seems to be more important to the system. Having 100 identical sentences scattered across the internet may not offer significant utility. A post that addresses 'The 5Ws' increases the likelihood of effective assistance compared to a generic 'Help' message. Future work includes addressing these technical challenges and exploring the integration of advanced models for enhanced performance.

Reference

[1] Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze. 2008. Introduction to Information Retrieval. Cambridge University Press.

[2] Cody Buntain, Richard McCreadie, Ian Soboroff. 2020. Incident Streams 2020: TRECIS in the Time of COVID-19. In *Proceedings of The Twenty-Ninth Text REtrieval Conference (TREC 2020),* Gaithersburg, MD, USA.

[3] Ian Soboroff. 2021. Overview of TREC 2021. In *Proceeding The Thirtieth Text REtrieval Conference(TREC 2021),* Gaithersburg, MD, USA.

[4] Congcong Wang, David Lillis. 2021. UCD-CS at TREC 2021 Incident Streams Track. In *Proceeding The Thirtieth Text REtrieval Conference(TREC 2021),* Gaithersburg, MD, USA.

[5] Jayr Pereira, Robson Fidalgo, Roberto Lotufo, Rodrigo Nogueira. 2022. Using Neural Reranking and GPT-3 for Social Media Disaster Content Summarization. In *Proceeding The Thirty-First Text REtrieval Conference (TREC 2022),* Gaithersburg, MD, USA.