Scientific Article Summarization: A Comparative Analysis of NLP Techniques

Mohammad Olid Ali Akash

Supervisor: Dr. Mohammad Ashrafuzzaman Khan

Department of Electrical and Computer Engineering | North South University

Abstract

This project analyzes various NLP techniques for summarizing scientific articles using models like LED, BART, and Zephyr, enhanced by advanced preprocessing and prompt engineering. Our unique approach integrates multiple state-of-the-art models and outperforms paid solutions like ChatGPT and Gemini. Offered for free, our summarizer aims to make high-quality academic summarization accessible to all researchers.

Motivation

Our project aims to streamline access to scientific knowledge using advanced NLP techniques. By simplifying complex articles, we empower researchers and students with concise, accessible summaries, enhancing learning and fostering innovation.

Dataset Details

- ✓ **Dataset Name:** Scientific Papers
- ✓ **Source:** Hugging Face Datasets
- ✓ **Dataset Summary:** The Scientific Papers dataset consists of two comprehensive sets of long and structured documents sourced from ArXiv and PubMed OpenAccess repositories.
- ✓ **Dataset Size:** The "arxiv" subset details are as follows:
  - Size of downloaded dataset files: 4.50 GB
  - Size of the generated dataset: 7.58 GB
  - Total amount of disk used: 12.09 GB

Methodology

- ❑ **Data Preprocessing:** Cleaned and pre-processed scientific articles using spaCy for lowercasing, punctuation handling, stop word removal and named entity recognition.
- ❑ **Model Training:** Fine-tuned transformer models like Falconsai, bart-large-cnn, zephyr-7b-beta and allenai/led-base-16384 on pre-processed data using Hugging Face's Seq2SeqTrainer. Employed techniques such as tokenization, attention masking, and global attention to enhance model performance.
- ❑ **Evaluation:** Evaluated model performance using metrics like ROUGE for summarization quality assessment. Compared generated summaries with reference summaries from datasets using ROUGE scores.
- ❑ **Prompt Engineering:** Utilized prompt engineering to tailor transformer models for specific tasks.

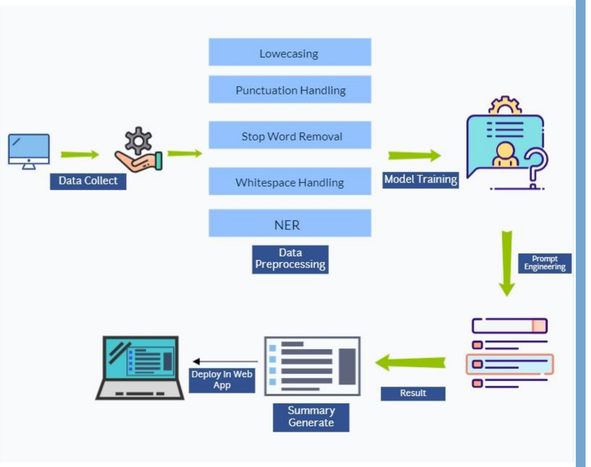
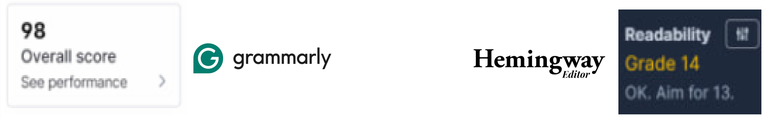


Figure Name: Full System Diagram

Result Analysis & Evaluation

allenai/led-large-16384: Anxiety affects the quality of life in those living with Parkinson's disease (PD) more so than depression. Anxiety affects cognition in PD patients with and without anxiety. Anxiety is a common and essential factor in developing executive functioning and working memory in PDS. The present study examined the relationship between anxiety and cognition in the central executive functioning of PD and the neuropsychological and neurophysiological domains of working memory and attentional systems of PDS.

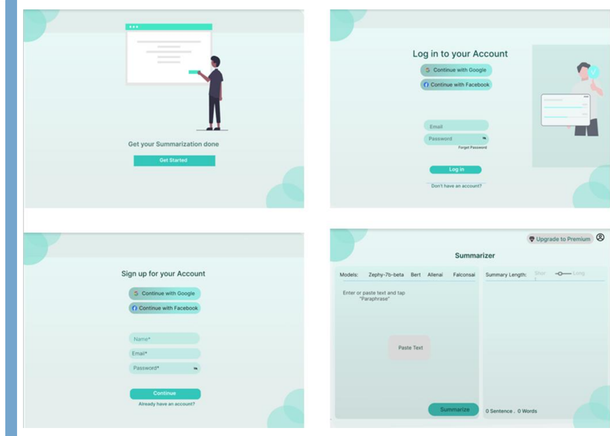
Gemini: This article investigates the relationship between anxiety and cognition in Parkinson's disease (PD). It finds that anxiety in PD patients is associated with worse performance in attentional set-shifting and working memory tasks. The article proposes that anxiety may disrupt compensatory mechanisms in the brain, leading to cognitive decline. It calls for future research to better understand the mechanisms and treatment of anxiety in PD.



ChatGPT: Anxiety significantly impacts cognition in Parkinson's disease (PD), particularly attentional set-shifting and working memory. PD patients with anxiety showed worse performance on tests assessing these functions compared to those without anxiety. Neurobiological changes, including dopamine depletion and hippocampal damage, may underlie these cognitive impairments. However, research on anxiety in PD remains limited. Future studies should explore neural correlates and treatment options to improve cognitive function and quality of life in PD patients with anxiety.

We compared the performance of three large language models (LLMs) on article summarization: allenai/led-large-16384,zephyr-7b-beta, Falconsai and bart-large-cnn .After fine-tuning our models with our datasets and employing prompt engineering techniques, we evaluated their summaries on various articles. We then compared these summaries to those generated by Gemini and ChatGPT using the same articles. To assess quality, we employed both automated methods. Our findings suggest that, in some cases, our fine-tuned models surpassed both Gemini and ChatGPT in terms of factual accuracy, clarity, engagement, delivery, and adherence to style guides.

Proposed Web Application



Conclusion

- ➢ Integrated models LED, BART, and Zephyr for robust summarization
- ➢ Outperformed paid solutions like ChatGPT and Gemini
- ➢ Offered free summarizer for wide accessibility
- ➢ Streamlined access to scientific knowledge
- ➢ Empowered researchers and students

Future Work

- ❖ Integration with Android and web applications
- ❖ OCR functionality for easy article upload
- ❖ Collaboration with educational and research institutions
- ❖ Free access to the web and Android app for institutions
- ❖ Continued prompt engineering for model optimization.
- ❖ Analysis of code and project for further enhancements