

ABSTRACT

RETRIEVAL-AUGMENTED GENERATIVE QUESTION ANSWERING SYSTEM AS MEDICINE-FOCUSED SEARCH ENGINE

by

Adika Farulia Zharfan
19/440441/PA/19071

The use of online means in various public services has become a common practice nowadays, including the health industry. While online public medicine information repository like Drugs.com stores an abundant amount of drug-related information, its search engine is not very optimized as it does not offer an answer to what the users are searching for. While studies in the topic of Question Answering (QA) have been trying to solve such issue in various fields by inventing a system that utilizes generative models and a relevant primary source of data to give answers instead of relevant articles for the user query, how the system performs using supplementary data is still need to be examined. As the medical field is known to have shortages on label information, this study aims to design a QA system using the retrieval-augmented generative approach in the subject of medicines.

The developed generative QA system used a patient drug review dataset as the proprietary data that was pre-processed at first and then passed to the Chroma vector database to supply the Large Language Model (LLM) with an additional context upon receiving a user query and performing similarity search. Initial BLEU evaluation identified Flan-T5-Base as the most optimal LLM with a score of 0.1004, leading to its implementation in the generative QA system. For findings comparison purposes, a basic generative QA system was created using all the assessed LLM.

A test using a set of 15 questions were done to the two generative QA systems. The results indicate improved accuracy in answering medicine-related queries and a reduction in the generation of hallucinated responses, especially for the case of Flan-T5-Base. This research addresses the need for optimized search capabilities in the health industry, specifically in overcoming label information shortages within the medical field.

Keyword: Natural Language Processing, Question Answering, Retrieval Augmented, Large Language Model, Chroma, Medicines, Drugs, BLEU metrics, Hugging Face