

LLM-based question answering for the infectious disease domain

Alexander Bondarenko,¹ Harrison Scells,¹ Martin Potthast,^{1,2} Adrian Viehweger³

¹Leipzig University, Germany

²ScaDS.AI, Germany

³Universitätsklinikum Leipzig, Germany

Introduction

Using inappropriate antibiotics can quickly lead to antimicrobial resistance and infections that are difficult to treat. However, the capacity for training expert medical personnel is very limited, especially in developing countries. To ameliorate this gap, we develop a personalized AI-supported tutor for antibiotic usage training, with a chatbot as one of its core components.

Methodology

Question: Which empiric therapy should be used for erysipelas in a patient with kidney failure and diabetes mellitus type 2?

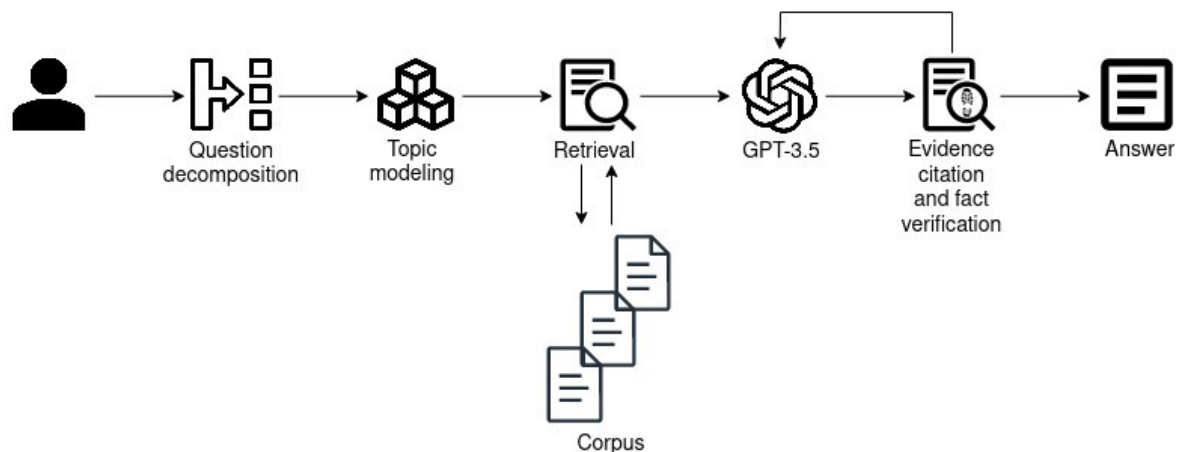


Figure: Overall chatbot architecture.

At the heart of the chatbot is a set of large language models (LLMs). However, LLMs are known to “hallucinate”, i.e., generate factually incorrect information. To tackle this, we use a retrieval augmented generation (RAG) approach that restricts the LLM response to the evidence context retrieved from a curated corpus of credible data. In our case, we collected an intentionally diverse set of media from guidelines and (review) articles to podcasts and blog posts across several levels of evidence quality.

Our retriever combines lexical retrieval models (e.g., BM25), re-ranking based on the question–document semantic similarity, and topical routing. The routing filters out retrieved documents whose modeled topic mismatches the question topic (e.g., “skin infection”).

To tackle complex questions (e.g., “Which empiric therapy should be used for erysipelas in a patient with kidney failure and diabetes mellitus type 2?”), we implement question decomposition that creates a set of shorter queries (e.g., “empiric therapy for erysipelas”, “therapy with kidney failure”, etc.).

Furthermore, we add the answer fact verification and source evidence citation modules. The fact verification module ensures that the generated answer’s statements are actually

supported by the retrieved evidence, and evidence citation allows the user to verify the statements by inspecting the evidence sources.

Since the development of the chatbot is ongoing, evaluation is part of future work.

Summary

We briefly described our ongoing work on developing an AI-supported chatbot that assists in the rational selection of antibiotic therapy. The chatbot architecture combines RAG with fact verification and evidence citation modules. The planned evaluation will reveal existing drawbacks and allow us to address them.