

Fact-Checking Health Claims Using Transformers and RAG Techniques

Abstract:

The proliferation of medical misinformation on digital platforms, particularly social media has emerged as a significant public health concern. The ability to accurately verify the information presented in such platforms is critical to preventing the spread of false or misleading health-related content. This project aims to develop an advanced system that leverages state-of-the-art natural language processing (NLP) techniques, specifically the BioBERT model, to verify medical claims through evidence-based fact-checking. The system begins by accepting user-provided textual input, which may contain medical terms or health-related claims. Using BioBERT, the system performs Named Entity Recognition (NER) to extract relevant medical entities, such as diseases, symptoms, treatments, and pharmaceutical products, from the input text.

Once the entities are identified, the system employs Retrieval-Augmented Generation (RAG) techniques to query multiple authoritative medical databases, including PubMed, the World Health Organization (WHO), and the Unified Medical Language System (UMLS). These databases provide reliable, up-to-date, and peer-reviewed medical information, allowing the system to cross-check the claims made in the input data. The RAG model is employed to retrieve evidence from these sources and assess the factual accuracy of the extracted entities, by comparing the retrieved data against the original claims.

The system's output is presented in a clear and user-friendly manner, categorizing the content into three possible outcomes: "Factual," indicating that the claim is supported by reliable evidence; "False," when the claim contradicts authoritative sources; and "Insufficient Evidence," when the claim lacks sufficient corroborative data or is not addressed in the databases. This real-time verification process helps users assess the validity of health-related claims and reduces the risk of misinformation dissemination.

By integrating the BioBERT model for accurate medical entity extraction and RAG for querying multiple sources, this system offers a robust solution for fact-checking health content. The project addresses a growing need for reliable tools to combat medical misinformation, ensuring that users can make informed decisions based on evidence-backed information. This system not only enhances the credibility of online health content but also contributes to broader efforts to improve public health literacy and reduce the impact of false medical claims in digital media.