

Harnessing LLMs for Generating Patient Discharge Reports

Norman Süsstrunk and **Albert Weichselbraun**
norman.suesstrunk@fhgr.ch

Abstract

Administrative processes in healthcare are important, but time-consuming and costly for all parties involved. The developed system aims to enhance the efficiency of documentation processes in psychiatric clinics using artificial intelligence (AI). The goal is to reduce the workload of qualified medical personnel, improve report quality, and reduce overall costs for psychiatric clinics. The project is being developed in collaboration with the Psychiatric Clinic Graubünden (PDGR). Technologically, the Report Recommender is based on harnessing Large Language Models (LLM's) combined with Multi-Document Summarization (MDS). This involves creating a summary from a collection of documents related to a specific topic. This process encounters challenges similar to summarising a long document, as the issue of limited context length becomes more pronounced with multiple documents. Understanding the relationships between these documents is crucial for an effective summary. One common strategy for MDS is concatenating all documents into one large text and applying techniques designed for single-document summarization. However, this approach requires the model to process very long sequences. Therefore, a two-stage process, similar to that used for long document summarization, is commonly employed (Liu* et al., 2018). State-of-the-art approaches also use hierarchical architectures or graph-based methods to capture inter-document relations (Liu & Lapata, 2019). At the same time, MDS approaches increasingly aim to utilise pre-trained encoder-decoder models such as BART, T5, or PEGASUS (Pasunuru et al., 2021). Further, identifying and eliminating redundant information while retaining relevant content is a complex task, particularly when dealing with large volumes of text. Extractive and abstractive summarization share inherent similarities; both are designed to extract the main information from an input natural language text. Current research (Zhang et al., 2023) indicates that the two tasks could mutually enhance each other. Extractive summarization selects key sentences and information from the original text, while abstractive summarization, with the assistance of large language models, rewrites and condenses information coherently. This approach aims to maintain the integrity of medical data while ensuring that the summary is concise and easy to understand. Therefore, a hybrid approach combining information extraction and abstractive summarization will be developed and evaluated.