Introduction

The integration of Large Language Models (LLMs) in healthcare is transforming clinical documentation by automating medical note generation and text summarization. As the reliance on Electronic Health Records (EHRs) grows, healthcare professionals face an overwhelming administrative burden that reduces the time available for direct patient care and increases clinician fatigue. LLMs, trained on extensive medical corpora, have the potential to streamline documentation by extracting relevant clinical details, summarizing lengthy records, and ensuring consistency in medical notes. These advancements can enhance workflow efficiency, reduce errors, and standardize medical documentation. However, despite their potential, the implementation of LLMs in real-world clinical settings remains complex due to challenges related to accuracy, interpretability, bias, and regulatory compliance. Ensuring that these models generate clinically accurate, unbiased, and interpretable medical notes requires further research, particularly in fine-tuning LLMs for domain-specific tasks, developing robust evaluation metrics, and addressing ethical considerations.

Motivation

Medical documentation is a time-consuming process that significantly impacts the efficiency of healthcare delivery. Studies indicate that clinicians spend nearly half of their working hours on documentation rather than direct patient interaction, leading to increased workload and physician burnout. Inefficiencies in note-taking also introduce risks of errors and inconsistencies in patient records, which can affect clinical decision-making and patient outcomes. LLMs offer a promising solution by automating documentation, summarizing key clinical details, and reducing the administrative burden on healthcare professionals. However, despite these benefits, several research gaps need to be addressed to ensure the safe and effective deployment of LLMs in medical settings. Challenges such as hallucinations, where models generate incorrect or misleading information, remain a significant concern. Additionally, domain-specific adaptation is required, as general-purpose LLMs may lack the specialized medical knowledge necessary for accurate documentation. Ethical concerns related to bias, fairness, and transparency must also be addressed to ensure that AI-generated medical notes do not introduce disparities in healthcare. Further research is needed to develop standardized evaluation frameworks, improve model reliability, and create interpretability mechanisms that allow clinicians to trust and validate AI-generated documentation. By addressing these challenges, this research aims to contribute to the advancement of LLMs in healthcare, ultimately improving documentation efficiency, reducing clinician workload, and enhancing patient care.