

先端科学技術研究科 修士論文要旨

所属研究室 (主指導教員)	自然言語処理学 (渡辺 太郎 (教授))		
学籍番号	2411066	提出日	令和 8年 1月 19日
学生氏名	尾崎 慎太郎		
論文題目	Understanding the Impact of Confidence in Retrieval Augmented Generation: A Case Study in the Medical Domain		
要旨			
Retrieval Augmented Generation (RAG) complements the knowledge of Large Language Models (LLMs) by leveraging external information to enhance response accuracy for queries. This approach is widely applied in several fields by taking advantage of injecting the most up-to-date information, and researchers are focusing on understanding and improving this aspect to unlock the full potential of RAG in such high-stakes applications. However, despite the potential of RAG to address these needs, the mechanisms behind the confidence levels of its outputs remain underexplored. Our study focuses on the impact of RAG, specifically examining whether RAG improves the confidence of LLM outputs in the medical domain. We conduct this analysis across various configurations and models. We evaluate confidence by treating the model's predicted probability as its output and calculating several evaluation metrics, which include the calibration error method, entropy, the best probability, and accuracy. Experimental results across multiple datasets confirmed that certain models possess the capability to judge for themselves whether an inserted document relates to the correct answer. These results suggest that evaluating models based on their output probabilities determines whether they function as generators in the RAG framework. Our approach allows us to evaluate whether the models handle retrieved documents.			