The Comparative Study on Recommendation Systems Based on Large Language Models

Yoshiho Kataoka

Large Language Models (LLMs), such as ChatGPT by OpenAI and Bard by Google, are deep learning models trained on large amounts of text data and can perform a variety of Natural Language Processing (NLP) tasks, such as sentence generation, classification, and summarization. Besides, there has been a growing need for recommendation systems that provide users with relevant information from the vast amount of information available. However, the conventional recommendation systems using deep learning can sometimes be overfitting to the training data, resulting in unsuccessful recommendations. LLMs have shown potential in solving these problems underlying the recommendation systems due to their strong capacities of domain generalization. In particular, LLMs can be adapted to specific tasks including recommendation tasks by devising the input which called prompt without fine-tuning, achieving a certain level of accuracy.

In this work, among the various recommendation tasks, we focus on the recommendation task of reranking a given list of candidate items. The research purpose is to investigate to what extent the different ways of formulating the prompts entered into the LLMs affect the precision of items' ranking. Specifically, we test an existing prompting method that relies on LLMs to re-rank documents in the field of Information Retrieval (IR), and a role-play prompting method that includes the description of role played by LLMs in the prompts. The experimental results show that these two methods lead to inferior ranking performance. A possible reason is that the prompt templates proposed in this work are redundant for LLMs. The performance could be improved by incorporating the strengths of the baseline prompts.

（Advisor : Hai-Tao Yu）