

QUESTION CONTEXT RETRIEVAL, BASED ON TOC LOOKUP IN QUESTION GENERATION AND TESTING SYSTEMS

Simeon Monov, Emre Myummyun, Nikolay Pavlov, Andrey Nikolov

Abstract. *Despite the advancements in large language models (LLM), question generation (QG) remains a difficult task to manage. Modern Generative LLMs can perform well on QG on pretrained data [1] but QG on domain-specific data requires pretraining and fine-tuning custom models on the domain-specific contexts and question-answer pairs [2]. Another approach is to select a smaller question context from a domain-specific document and provide it as part of the prompt to the LLM. Selecting the correct context is challenging especially when working with large documents.*

In this paper we propose a method of retrieval and selection of question context, based on the document table of contents (TOC) using LLM. The method is very effective on tasks, where a sequence of questions must be generated, so that each question depends on the previous ones. The method can be applied in automated testing for job interviews or student examinations.

Key words: Question context, question generation, LLM, context retrieval.

Acknowledgments

This study is supported by the project FP23-FMI-002 “Intelligent software tools and applications in research in mathematics, informatics, and teaching pedagogy” at the Plovdiv University “Paisii Hilendarski”.

References

- [1] X. Yuan, T. Wang, Y. Wang, E. Fine, R. Abdelghani, P. Lucas, H. Sauzéon, P. Oudeyer, Selecting better samples from pre-trained LLMs: A case study on question generation, arXiv preprint arXiv:2209.11000, 2022 Sep 22.
- [2] A. Hasan, M. Ehsan, K. Shahnoor, S. Tasneem, Automatic question & answer generation using generative Large Language Model (LLM) (Doctoral dissertation, Brac University).

Simeon Monov¹, Emre Myummyun², Nikolay Pavlov³, Andrey Nikolov⁴,

^{1,2,3,4} Paisii Hilendarski University of Plovdiv,

Faculty of Mathematics and Informatics,

236 Bulgaria Blvd., 4003 Plovdiv, Bulgaria

Corresponding author: smonov@uni-plovdiv.bg