Bachelor Thesis Topic

Automatic Derivation Of Prompt Templates (ADOPT)

Motivation and Background

Large language models (LLMs), such as ChatGPT, have rapidly gained popularity due to their ability to perform diverse natural language processing tasks effectively. However, crafting precise and efficient prompts for these models remains a significant challenge, often requiring substantial manual effort [1]. This thesis aims to address this challenge by developing an automated system capable of generating reusable prompt templates. Given an example prompt, the system will separate its structural template from its content, ensuring that recombining these two components reproduces the original prompt accurately. Once successfully extracted, the generated templates can be efficiently applied to new topics, maintaining consistent structure and reducing manual prompt engineering efforts.

Goals

The goal of this thesis is to:

1. **Implementation**. Implement a basic pipeline that takes a pair of content and the desired output and returns a prompt template that -- combined with the content -- would have led to that specific output if given to an LLM.
2. **Technique.** Develop and evaluate different methods to use in the pipeline.
3. **Evaluation dataset.** Build a small dataset of prompt and desired output pairs from a suitable domain.
4. **Evaluation of Accuracy.** Assess how well the generated templates work on new samples.

Research Type

Theoretical Aspects:	**∗∗∗
Industrial Relevance:	****∗
Implementation	*****

Prerequisite

The student should be enrolled in the bachelor of computer science program, and has completed the required course modules to start a bachelor thesis.

Skills required

We are seeking a candidate with **strong Python programming skills** and practical knowledge of large language models (LLMs), including hands-on experience. Completed courses in software engineering-II or machine learning research is a plus, as is the ability to communicate results effectively.

Contacts

enrico.philip.ahlers@hu-berlin.de
Software Engineering, Institut für Informatik, Humboldt-Universität zu Berlin

References

[1]: Sahoo, Pranab, et al. "A systematic survey of prompt engineering in large language models: Techniques and applications." arXiv preprint arXiv:2402.07927 (2024).