

Prompt Revision for LLMs

LLMs are awesome, but finding (or the art of designing) a satisfactory prompt is not easy [1]. For example consider the following prompts and their corresponding ChatGPT responses:

Prompt 1: Write a poem.

Response 1:

In the quiet of the night,
A solitary moon takes flight,
Its silver beams caress the land,
A gentle touch from a celestial hand.

Prompt 2: Write a haiku about the changing seasons.

Response 2:

Autumn leaves fall slow,
Winter's breath chills,
spring buds grow,
Summer sun aglow.

The second prompt achieves a completely different poem than the first through the technique of a more precise formulation. Techniques out of this area are called prompt engineering which is the process of designing and refining input queries to large language models. Examples of such prompting techniques are Least-To-Most Prompting or Chain-Of-Thought Prompting [2]. They could be broadly categorized in three categories: "Basic", "Advanced" and "Scientific" [2].

As there are many prompt engineering methods, it is not possible to try out all methods for every prompt. This work therefore aims to systematically compare as many methods as possible using existing benchmarks (e.g. BIG-bench from Google).

Instruction tuning is a related technique used to enhance the output of LLMs, which uses *Instruction-Response* pairs for fine-tuning an LLM to better adhere to the users' instructions and objectives [3]. The proposed work will explore instruction tuning besides prompt engineering as technique to make the output of LLMs more responsive and aligned with user instructions.

The thesis can be written in English or German. Familiarity with Python is required.

Related literature

- [1] Sahoo et al. 'A Systematic Survey of Prompt Engineering in Large Language Models: Techniques and Applications'. In: arXiv (2024), <https://arxiv.org/abs/2402.07927>
- [2] Chen et al. 'Unleashing the potential of prompt engineering in Large Language Models: a comprehensive review'. In: arXiv (2023), <https://arxiv.org/abs/2310.14735>
- [3] Zhang et al. "Instruction Tuning for Large Language Models: A Survey". In: arXiv (2024), <https://arxiv.org/abs/2308.10792>

Interested? @mail to fliedeker@techfak.uni-bielefeld.de