

УДК 004.8

АНАЛИЗ ЗАВИСИМОСТИ ОТВЕТА БОЛЬШОЙ ЯЗЫКОВОЙ МОДЕЛИ ОТ ФОРМАТА ОТПРАВЛЕННОГО ЕЙ ПРОМПТА

Иванов В. С. (Университет ИТМО), Насыров Н.Ф. (Университет ИТМО)

Научный руководитель - кандидат технических наук, доцент Федоров Д.А. (Университет ИТМО)

Введение. В работе с большими языковыми моделями (Large Language Models, далее - LLM) большую роль играют отправленные пользователем промпты (с англ. “prompts” - запросы, подсказки). Каждый промпт обладает такой комплексной характеристикой, как хрупкость. Она обозначает зависимость качества полученного результата от характеристик отправленного промпта. Для того, чтобы пользователь получил от LLM необходимый ему результат, ему необходимо решить такую задачу, как правильное составление промпта. Промпт-инжиниринг запросов включает в себя понимание нюансов использования LLM и тщательный анализ влияния каждого элемента промпта на конечный результат. Значение формулировки промпта очень велико.

Существуют различные методы улучшения промптов, опирающиеся на конкретный, необходимый пользователю результат, однако, если провести анализ поведения LLM, применяя различные техники инжиниринга промптов и проводя калибровку элементов каждого промпта, можно получить комплексное решение сведения хрупкости до минимума [1].

Основная часть. Для улучшения качества полученного от большой языковой модели результата и сведения хрупкости промптов до минимума необходимо проанализировать ее ответ в зависимости от изменяемых параметров, в т. ч. изменение обозначения входных данных и индикаторов выхода, расширение контекстного окна (контекстная калибровка), создание герменевтического круга, наличие обучающих данных и т. п. Оптимальной тактикой решения поставленной проблемы является комплексное решение, включающее в себя сразу несколько методов уменьшения хрупкости промптов [2]. При этом в исследовании не рассматривалось изменение гиперпараметров самой LLM.

Для проведения анализа влияния характеристик промптов на получаемый результат, была выбрана большая языковая модель Gigachat Pro. В основе GigaChat лежит нейросетевой ансамбль NeONKA (NEural Omnimodal Network with Knowledge-Awareness), включающий в себя модели ruGPT-3.5 с 13 млрд параметров, Kandinsky 2.1, ruCLIP и FRED-T5 [3]. В качестве предметной области рассматривается задача определения времени в предложении на английском языке. Извлеченная информация об используемом времени в англоязычном предложении может использоваться в процессе автоматизированного составления индивидуальных заданий для обучающихся, генерации тестовых заданий и т.д. LLM в ряде случаев не могла определить времена (tenses), кроме Past Simple, Present Continuous и Future Simple. Также были зафиксированы ошибочные ответы большой языковой модели на запрос определения времени в сложных предложениях (двусоставные, с прямой речью, вопросительные и т.д.). Таким образом, была сформулирована следующая задача: “определение времени в предложении на английском языке” с типом промпта “multilabel-классификация”.

В качестве датасета использовался набор из 200 заранее размеченных сложносоставных предложений на английском языке. Под корректным ответом понималось указание всех времен, использующихся в предложении. Для анализа хрупкости промптов использовались следующие техники промпт-инжиниринга:

1. Присвоение модели конкретной роли - задание LLM роли позволяет сделать промпт более контекстуальным и специфическим для конкретной задачи.

Промпт: Представь, что ты лингвист. Определи все verb tenses в предложении "She plays the piano beautifully, and she will perform at the concert next week" . Обрати внимание, что в одном предложении может быть 1 или 2 грамматических основ.

Выдай только ответ - 1 время или 2 времени на английском языке в этом предложении без объяснений правил выбора времени.

Корректных ответов - 72%

2. Few-Shot Prompting (FSP) - техника, заключающаяся в расширении контекстного окна языковой модели введением герменевтического круга [2].

Промпт: Определи время (tense) в предложении

“She said, “I study French every Monday” => Past Simple, Present Simple

...

“She plays the piano beautifully, and she will perform at the concert next week” =>

Корректных ответов - 78%

3. Tree of Thoughts (ToT) - подход позволяющий LLM самостоятельно оценивать прогресс, достигнутый промежуточными мыслями в решении проблемы, посредством целенаправленного процесса рассуждений. [4]

Промпт: Представьте, что на этот вопрос отвечают три разных лингвиста.

...

Если какой-либо эксперт понимает, что он неправ в какой-либо момент, он уходит.

Их задача - определить время (tense) в предложении “She plays the piano beautifully, and she will perform at the concert next week” Ответ - tense на английском языке.

Корректных ответов - 82%

4. Chain-of-Thought (CoT) Prompting - техника промпт-инжиниринга, обеспечивающая модели возможность создания логической цепочки с помощью промежуточных этапов рассуждения. [5]

Промпт: “I was doing my work all day long” - слово “was” указывает на то, что действие происходило в прошлом. Также, слово “doing” - продолженная форма глагола “to do” указывает на то, что действие было продолжительным => Ответ: Past Continuous.

...

“She plays the piano beautifully, and she will perform at the concert next week” =>

Корректных ответов - 85%

5. Применение нескольких приемов одновременно с целью повышения эффективности работы модели.

Использование комплексного промпта, сочетающего в себе вышеперечисленные методы, позволяет получить результат в 94%

В ходе исследования было выявлено, что:

1. рассматриваемая LLM в первую очередь опирается на контекст, заданный пользователем, а не на вложенные в саму модель данные, что подчеркивает важность грамотного составления и расширения контекстного окна
2. LLM способна определять “флаги” конкретного времени английского языка, однако часто неверно истолковывает их наличие в предложении
3. Лучший шаблон задания индикатора выхода это (Answer - Result)

Выводы. Работа с LLM требует тщательного составления промптов для минимизации их хрупкости. Анализ поведения больших языковых моделей при использовании различных техник промпт-инжиниринга позволяет улучшить процесс анализа текста и повысить качество конечного результата. Важно учитывать характеристики промптов и применять комплексный подход к их улучшению. В частности, промпт-инжиниринг с использованием метода Chain-of-Thought (CoT) оказался наиболее полезным, обеспечивая модели возможность создания логической цепочки рассуждений и значительно повышая качество анализа текста. Однако важно отметить, что комбинация различных приемов играет не меньшую роль. Например, совмещение методов CoT, Few-Shot Prompting и ToT может

улучшить не только точность, но и интерпретируемость полученных результатов, делая модель более эффективной в решении задач.

Список использованных источников:

1. Calibrate Before Use: Improving Few-Shot Performance of Language Models // arXiv preprint arXiv: 2102.09690v2 [cs.CL] 10 June 2021
2. Henrickson, L., Meroño-Peñuela, A. Prompting meaning: a hermeneutic approach to optimizing prompt engineering with ChatGPT. AI & Soc (2023) - С. 1-6.
3. Документация GigaChat. Возможности GigaChat [Электронный ресурс] - 2024. - URL: <https://developers.sber.ru/docs/ru/gigachat/about> (дата обращения 17.02.2024)
4. Prompt Engineering Guide [Электронный ресурс]. - 2024. - URL: <https://www.promptingguide.ai/> (дата обращения 15.02.2024)
5. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models // arXiv preprint arXiv: arXiv:2201.11903v6 [cs.CL] 10 Jan 2023

Иванов В. С.. (автор)

Подпись

Насыров Н.Ф. (автор)

Подпись

Федоров Д.А. (научный руководитель)

Подпись