

A. Zinovenko
А.Д. Зиновенко
БрГУ им. А.С. Пушкина (Брест)
Научный руководитель М.П. Концевой

MODELLIERUNG VON BENUTZERANFRAGEN IN GROSSEN SPRACHMODELLEN

Моделирование пользовательских запросов к большим языковым моделям

В современную практику языкового образования стремительно входят большие языковые модели (Large Language Models, LLM) — системы машинного обучения, использующие искусственные нейронные сети для решения задач автоматической обработки языка.

Большие языковые модели доступны для учащихся и преподавателей посредством чат-ботов, пользовательский интерфейс которых организован в виде диалоговых сервисов (YandexGPT2, GigaChat и др.). В диалогах с большими языковыми моделями можно получить не только содержательные ответы на вопросы по любому материалу, но и продуктивные результаты разноплановой речевой деятельности.

Качество получаемых от LLM ответов во многом определяется эффективностью пользовательских запросов (подсказок, промтов, промптов, prompts) на которые большие языковые модели дают ответы. Запросы (подсказки, промты, промпты, prompts) представляют собой фразы, или синтаксически неоформленные наборы (последовательности), с помощью которых пользователь обращается к LLM, описывая задачу и запуская генерацию текста, который будет ответом (результатом выполнения задачи). От формулировки запроса, его лексического наполнения и грамматического оформления зависит качество ответа со стороны нейросети. Здесь действует принцип «garbage in, garbage out» — на плохо сформулированный запрос вы получите невнятный ответ. Ответ генеративной нейросети может существенно различаться даже в зависимости от использованной в запросе последовательности слов в простом их перечне.

Значимость пользовательских запросов для построения эффективного взаимодействия с LLM привело к появлению Prompt engineering, новой области профессиональной деятельности. Суть Prompt engineering заключается в разработке текстовых запросов на естественном языке (prompts), которые будут оптимальным образом интерпретироваться и выполняться генеративными нейросетями. Практическое решение проблемы эффективности запросов к LLM для обычных пользователей реализовано посредством предложения на

различных сетевых ресурсах множества готовых шаблонов и образцов; инструкций по составлению результативных запросов; сервисов автогенерации запросов. Многие из таких ресурсов несомненно являются полезными и нужными инструментами в поддержании диалогов и большими языковыми моделями на приемлемом для пользователя уровне. Вместе с тем, даже грамотное использование лучших из таких инструментов никак не гарантирует эффективности в диалоге с LLM, более того, частотно значимо сопровождается полными неудачами в построении такого диалога. Таким образом, подобные инструменты не решают проблему.

Представляется, что настоящее решение проблемы повышения эффективности пользовательских запросов к LLM требует их теоретического осмысления на более высоких уровнях обобщения, соответствующих лингводидактическому контексту. Например, неправильно составленный пользовательский запрос может стать причиной полного отказа LLM обсуждать затронутую в нем проблематику. Часто причина такого отказа заключается в том, что LLM просто не может правильно интерпретировать пользовательский запрос. Нередко, даже самая незначительная, по мнению пользователя, корректировка запроса к той же самой LLM позволяет в ряде случаев получить от неё блестящий ответ. В лингводидактическом контексте это может быть объяснено тем, что нейросеть работает с языком и речевым материалом совершенно иначе, чем человек. Язык в машинном представлении нечто кардинальное иное, чем язык человека. Взаимодействие человека (пользователя) и LLM (генеративной искусственной нейросети) может быть понято в парадигме межъязыковой коммуникации.

E. Kozich

Е.С. Козич

БрГУ им. А.С. Пушкина (Брест)

Научный руководитель М.П. Концевой

VERWENDUNG DER KORPUSLEXIKOGRAPHIE IN EINEM DEUTSCHUNTERRICHT

Использование корпусной лексикографии в преподавании немецкого языка

Der Einsatz der digitalen Lexikographie erfreut sich in der Lehre immer größerer Beliebtheit. Die Frage ist, wie man die Online-Wörterbücher und Elektronische nationale Korpora im Studium benutzen kann? Mit Hilfe der Korpora kann man: Abweichungen von der Standardlautung notieren; Mehrdeutigkeit von