

PROMPT ENGINEERING – ТЕХНОЛОГИЯ ОБЩЕНИЯ С НЕЙРОСЕТЬЮ

Брехова А.Г.

Томский политехнический университет, ИШИТР, 8К11, e-mail: agb17@tpu.ru

Введение

В 2021-2022 гг. произошел «бум» систем для генерации изображений, текстов и звуков, основанных на технологиях машинного обучения и нейронных сетей. Генеративные нейросети (далее - ГН) уже стали помощниками для людей, работающих в таких сферах как IT, медиа, продакшн и дизайн.

30 ноября 2022 года вышел ChatGPT [1] от компании OpenAI на базе GPT-3.5. ChatGPT – это нейросеть, имеющая **диалоговый** интерфейс (Рис.2) и отвечающая на запросы пользователей в реальном времени. Во многом благодаря этой сети, а также Dall-e [2] от OpenAI и другим ей подобным, что генерируют различные виды информации по запросу пользователя, появилось такое понятие как *“prompt engineering”* (от англ. «prompt» - запрос). О том, что оно собой представляет и пойдет речь в данной статье.

Промпт инжиниринг – профессия будущего

На данный момент существует несколько видов ГН, самые распространенные из них:

1. Text-to-text – генерация текста по текстовому запросу;
2. Text-to-image/video/audio/3D – генерация медийного (кроме текста) контента по текстовому описанию;
3. Text-to-code – генерация кода по текстовым инструкциям.

Перед генеративной нейросетью стоит задача как можно качественнее сгенерировать нужный контент – будь то код, текст или изображение – который отражает промпт, сделанный пользователем. Отсюда и происходит новый, появившийся в сетях термин: промпт инжиниринг.

Промпт-инженер – человек, который знает, что нужно написать в строку ввода, чтобы машинный интеллект сгенерировал приближенное к этому значение, соответствующее виду нейросети. Оказывается, сделать это не так-то просто, так как существуют определенные ограничения по возможностям нейросети и по базе данных, на которой она обучена.

Предложенные вниманию читателя примеры показывают наиболее распространенные «артефакты» (на жаргоне промпт-инженеров – «ошибки») генеративных нейронных сетей:

- 1) Нечитабельный текст на картинках. Данный недостаток, однако, смогли обойти создатели нейросети «DeepFloyd», сделав текст на картинках в точности повторяющий запрос пользователя.
- 2) Несоответствие фактов действительности

Приведем пример для нейросети «text-to-text», ChatGPT. Данная сеть может неправильно указывать даты событий или некоторые значения слов, и при этом говорить о них совершенно уверенно.

У нейросетей text-to-image или image-to-image часто прослеживается следующая тенденция: при генерации изображений они пренебрегают «мелочами», которые на самом деле довольно существенны. Показательным примером является количество пальцев на руках (рис. 3(а)) и искажение анатомических особенностей живых организмов (рис. 3(б)).

«Нереальность» изображений придает им эффект «во сне» или «утекающих часов» – в честь известного художника и названа сеть Dall-e и ее русский аналог от Sber AI: Ru-Dall-e [6].

Чтобы обойти вышеперечисленные и другие ограничения, промпт-инженер должен:

- знать, на каких данных обучена нейросеть;
- понимать специфику конкретной нейросети, ее лучшее применение;
- давать более подробные описания своих запросов;
- приводить примеры для нейросети и др.

Уже начали выпускаться специальные гайды – инструкции – для тех, кто хочет работать с нейросетями и рассматривает их для постоянного использования. Примером таких гайдов является статья от канадской компании Cohere-AI, где описывается принцип работы нейросетей и то, как лучше всего с ними взаимодействовать [7].

Появилось также направление автоматического промпт инжиниринга (АПИ). Так как количество синонимичных по смыслу промптов неограниченное количество, и на обработку их всех нужно время, группа программистов создала автоматический генератор промптов, который определяет, какой лучше

вариант 1-го промпта подать на вход нейросети [8]. Промпты генерируются компьютером, на основе данных об их «понятности» для ИИ отбираются лучшие, затем получается конечный результат.

Пример работы АПИ представлен на рис. 1. В качестве вариантов было выбрано (перевод с английского): «Напиши антоним к этому слову», «Дай антоним данного слова», «Напиши слово, противоположное данному» .

write the antonym of the word.	-0.26	✓
give the antonym of the word provided.	-0.28	✓
...	...	
reverse the input.	-0.86	✗
to reverse the order of the letters	-1.08	✗
write the opposite of the word given.	-0.16	★
...	...	
list antonyms for the given word.	-0.39	

Рис. 1. Варианты промптов и их логарифмическая вероятность понимания ГС

За этим и многими другими аспектами промпт инжиниринга кроется весьма практичный смысл. Чем лучше ГС понимает запрос, тем:

- лучше качество обработки запроса;
- выше ее скорость;
- меньше количество потраченных токенов (единица измерения работы нейросети, оплачиваемая (при необходимости) ее пользователем)

Заключение

Несмотря на то, что смысл генеративных сетей – помогать людям решать их задачи, они не смогут автономно работать, поэтому нужны промпт инженеры, выступающие связующим звеном между идеей и ее реализацией с помощью ГС. Такая профессия, вполне возможно, скоро войдет в список востребованных в сфере IT.

Список использованных источников

1. ГНС ChatGPT. URL: <https://openai.com/blog/chatgpt/>
2. ГНС Dall-e от OpenAI. URL: <https://openai.com/dall-e-2/>
3. ГНС Stable Diffusion. URL: <https://stablediffusionweb.com/>
4. Сайт компании по разработке ИИ - Stability AI. URL: <https://stability.ai/>
5. Генеративная нейронная сеть QQ. URL: h5.tu.qq.com/web/ai-2d/cartoon/index
6. Генеративная нейронная сеть Ru-Dalle. URL: <https://rudalle.ru/>
7. Prompt Engineering. URL: <https://docs.cohere.ai/docs/prompt-engineering>
8. APE: Automatic Prompt Engineer. URL: <https://sites.google.com/view/automatic-prompt-engineer?pli=1>
9. Co-pilot. URL: <https://github.com/features/copilot/>