

«Идеальный партнёр по мозговому штурму»: московский IT-евангелист о прошлом и настоящем нейросетей, омнимодальности и о том, как искусственный интеллект изменит нашу жизнь в будущем

О правилах и особенностях взаимодействия с ИИ рассказал доцент Московского авиационного института (МАИ) и НИУ ВШЭ Дмитрий Сошников.



Фото: Пресс-служба МАИ / Личный архив

Ошибочность, поверхностность и безэмоциональность производимого нейросетями контента во многом связаны с неправильными действиями человека-оператора. Чтобы получать от ИИ качественные результаты, важно научиться правильно с ним работать — ни одна нейросеть не способна сгенерировать хорошую картинку или текст по одному-единственному запросу. Освоить искусство работы с нейросетями позволяет такая дисциплина, как промпт-инжиниринг. О правилах и особенностях взаимодействия с ИИ рассказал доцент Московского авиационного института (МАИ) и НИУ ВШЭ Дмитрий Сошников. Он объяснил, какие «врождённые» ограничения мешают нейросетям качественно работать без участия человека, а также поделился прогнозом относительно будущего нейросетевых технологий. Сошников не исключает, что в отдалённом будущем ИИ сможет даже читать и формулировать мысли человека.

— Недавно глава Nvidia Дженсен Хуанг заявил, что в будущем каждый специалист будет прибегать к помощи ИИ-ассистента вне зависимости от сферы деятельности. Генеративные нейросети действительно произвели фурор, однако сейчас уже ясно, что ИИ не способен заменить человека. Написанные нейросетями тексты поверхностны и нуждаются в проверке, а сгенерированные картинки лишены эмоций и содержат ошибки. Какую в итоге роль будут играть нейросети?

— В действительности нейросети могут создавать качественный и вызывающий у людей эмоциональный отклик контент. Просто для этого нужно, чтобы человек принимал активное участие в процессе: сама по себе нейросеть не может испытывать эмоции, отделять правду от вымысла. Поэтому человек должен выступать в роли верификатора, а также задавать ИИ правильные запросы для генерации. Вообще, нейросети способны дать видимый эффект во всех сферах человеческой деятельности, но только если ими правильно пользоваться. Ведь часто пользователь ожидает, что нейросеть сделает всё за него, но это невозможно. Я часто говорю своим знакомым: «Наступает эпоха искусственного интеллекта, возьмите ChatGPT и примените его в своей работе». Они пробуют, а потом говорят: «Мы спросили, как увеличить продажи, а он ответил: „Снизьте цену“. Но это мы и так знали».

На мой взгляд, проблема состоит в том, что люди неправильно используют нейросети: на очевидный вопрос неизбежно получаешь очевидный ответ. Нужно ставить перед нейросетью более сложные задачи, тогда она будет пытаться их решать.

— Другими словами, нужен правильный метод работы с нейросетями?

— Верно. Кроме того, у нейросетей есть фундаментальное ограничение, которое приводит к тому, что они часто пишут «воду». Нейросеть не может думать как человек, она просто генерирует правдоподобный текст. Однако в партнёрстве с человеком она способна дать хороший эффект. Пользователь должен не только правильно составлять запросы, но и выступать в качестве критика, отбирать нужный ему результат и направлять работу ИИ. Тогда нейросеть действительно станет инструментом для решения нестандартных задач. Например, можно использовать нейросеть для мозгового штурма — просить её генерировать сразу много идей, а потом отбирать лучшие. А дальше генерировать уже отдельные составляющие этих идей. Нейросеть способна создавать текст или изображение намного быстрее, чем любой копирайтер или дизайнер, и при правильном подходе её использование может ускорить работу в разы. Большинство людей не могут производить настолько же разнообразные и разноплановые идеи, как нейросеть. Это и понятно: мозг человека, как правило, идёт по проторённой дорожке, оценить проблему со всех сторон ему сложно. Именно поэтому для мозгового штурма привлекают разных людей. А у нейросети нет таких ограничений, поэтому она является идеальным партнёром по мозговому штурму, способным заменить сразу нескольких людей.

— Сгенерированные нейросетью тексты часто содержат ошибки и ложную информацию, сгенерированные картинки тоже нуждаются в проверке. В какой области нейросетевые технологии достигли наилучших результатов, по вашему мнению?

— Сложности с текстовым контентом связаны с тем, что нейросеть — это некая вероятностная модель. Она генерирует наиболее правдоподобный текст, который будет очень похож на то, что она видела в процессе обучения, но при этом про истинность этого текста речь не идёт вообще. Очень часто случаются так называемые галлюцинации, когда нейросеть пишет что-то такое, что не имеет отношения к реальности. Часто это связано с ещё одной базовой особенностью генеративных нейросетей: поскольку они учились на текстах, которые в большей степени носят позитивный характер (простое изложение материала), им проще давать утвердительные ответы. Это первое. А второе — у нейросети нет своего мнения, поэтому, если ей сказать: «Приведи десять причин, почему гомеопатия работает», она их приведёт, а если сказать: «Приведи десять причин, почему она не работает», она тоже их приведёт. Просто нейросеть училась на всём, какого-то своего мнения у неё может и не быть. Например, некоторое время назад у меня был любимый вопрос к нейросети: «Объясни, почему число „пи“ больше пяти». ChatGPT приводил обоснования в пользу этого и при этом не говорил: «Нет, это утверждение ошибочно». Поэтому нейросети не следует напрямую использовать для генерации контента в социальных медиа. Да и вообще, пытаться с помощью ИИ создавать сразу, за один запрос, готовый контент — не лучшая идея. В сфере образования, когда школьники пишут целые сочинения с помощью генеративного ИИ, такое может привести к тому, что дети разучатся

сами писать и формулировать мысли, а ведь это крайне важно для развития мышления. По сути, наше мышление и есть процесс оформления мыслей в текст, речь.

— Изображения и тексты у всех на слуху, а как насчёт видео?

— В генерации видео мы находимся на такой промежуточной стадии, когда результат уже очень убедителен, но ещё нуждается в доработке. Судя по тем моделям, которые доступны, заметно, что видео сгенерировано, и оно, как правило, короткое. Для социальных медиа это не страшно, потому что они в основном и используют нарезку каких-то коротких фрагментов. Но фильм, например, силами искусственного интеллекта ещё снять нельзя, до этого ещё должно пройти некоторое время. Однако видеотрекеры, которые выкладывают после использования тестовой версии последней нашумевшей модели OpenAI Sora, выглядят убедительно. Есть ожидание, что создание видеоконтента при помощи искусственного интеллекта тоже станет повседневной реальностью.

— На ваш взгляд, нейросети — временное увлечение людей или они останутся с нами надолго?

— Это, безусловно, не временный хайп, поскольку они позволяют автоматизировать большой пласт задач, которые раньше считались исключительно прерогативой человека. Но нейросети не заменят человека в обозримой перспективе. И чтобы этого не произошло в принципе, нужно не выпускать инициативу из своих рук: например, уже сегодня очень важно учиться применять нейросети в своей работе, обучаться методам промпт-инжиниринга.

— Поясните, пожалуйста, о чем идёт речь.

— Этот модный термин означает умение общаться с нейросетями и добиваться от них наилучших результатов. Начиная с умения чётко ставить задачу и заканчивая такими тонкостями, как применение особых, «магических» слов, которые сильно улучшают качество контента. Например, Canon EOS 5D — это просто название фотоаппарата, но при использовании его при запросе к нейросетям можно существенно повысить качество картинок, потому что нейросеть видела много хороших студийных фотоизображений с такой подписью. Соответственно, когда мы используем это «магическое» слово, нейросеть понимает, на что примерно ориентироваться. Кстати, если кто-то хочет познакомиться с промпт-инжинирингом подробнее, а заодно и понять, как устроены генеративные нейросети, — приглашаю всех на открытый курс по промпт-инжинирингу, который мы записали с коллегой из Школы дизайна НИУ ВШЭ.

— Весной в Госдуме обсуждалась разработка закона, который обяжет маркировать созданный нейросетями контент. А есть ли способы отличить его от созданного человеком?

— Есть, но лишь с некоторой вероятностью. Например, недавно Сбер представил свой детектор, который позволяет отличать сгенерированный контент от созданного человеком. Как рассказали в компании, анализ разных социальных медиа, крупных новостных изданий показал, что в среднем 6% опубликованного контента сгенерировано ИИ. Для некоторых изданий эта доля достигает около 30%. В каком-то смысле отличить такой контент можно и без специальной программы, просто по частотному словарю. Оказывается, что генеративные сети имеют предпочтения к использованию определённых слов. ChatGPT почему-то очень любит слово delve (англ. «углублять»): частота его использования у нейросети больше, чем у человека. Когда эксперты проанализировали встречаемость слова delve в медицинских статьях за 2023 год, то увидели просто чудовищный всплеск. Это говорит о том, что не только журналисты, но и учёные начинают активно пользоваться нейросетями.

— В каком направлении сейчас развиваются нейросетевые технологии? Какие в этой области есть передовые разработки?

— Думаю, главная тенденция — создание омнимодального ИИ, то есть таких нейросетей, которые смогут воспринимать одновременно речь с учётом всех интонаций, текст, изображения и так далее. Почему это важно? Приведу такой пример. Если искусственный интеллект распознаёт только звук в речи и потом синтезирует ответ, то при этом теряются все эмоции. Человек может произносить что-то с эмоциональной окраской, с подтекстом, с иронией, и все эти интонации пропадают. Если мы берём омнимодальный искусственный интеллект, то он воспринимает всю речь целиком, включая эмоции, и может на неё соответствующим образом реагировать. На последней презентации OpenAI продемонстрировали, как работает их новейший омнимодальный ассистент, и он на эмоциональную речь отвечает эмоционально. Это очень близкий к человеческому уровень владения каналами поступления информации, что открывает перед пользователями большие возможности.

— Каковы ваши прогнозы на более отдалённое будущее?

— В принципе, текущего уровня развития генеративного ИИ уже достаточно, чтобы совершить революцию в сфере искусственного интеллекта. Но давайте попробуем заглянуть ещё дальше. Искусственный интеллект при генерации изображения оперирует многомерными пространственными образами. Раз компьютерная модель способна рисовать картинки по вербальному описанию, значит, мы уже сгенерировали некое виртуальное пространство, в котором каждая точка соответствует какому-то изображению или фрагменту текста. Следующим этапом может стать, например, считывание сигналов активности мозга человека с помощью функциональной МРТ и их соотнесение с заложенным в нейросети пространством смыслов и образов. Таким образом, мы сможем читать мысли человека. Такие исследования сейчас ведутся — правда, пока не очень успешно, но какие-то образы извлекать из сознания человека всё же удаётся. Возможно, со временем мы научим нейросети генерировать текст за человека по его мыслям. Учитывая словарный и понятийный запас ИИ, не исключено, что это позволит формулировать мысли даже точнее, чем это могут делать люди. Такие технологии станут новым инструментом коммуникации. Однако пока что это перспективы очень отдалённого будущего.

Материал подготовлен при поддержке Минобрнауки России.