

ГЕНЕРАТИВНАЯ МОДЕЛЬ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Мухамадиева Кибриё Баходировна

Соискатель кафедры «Компьютерный инжиниринг» Ташкентского
университета прикладных наук, г. Ташкент. Узбекистан

mkb78@mail.ru

АННОТАЦИЯ

Закрытые системы современного искусственного интеллекта, похоже, не приведут к созданию разумных машин в ближайшем будущем. Что необходимо, так это открытые системы с нелинейными свойствами, чтобы создать интересные свойства для каркаса искусственного разума. Это область генеративного искусственного интеллекта, в которой реализуются механизмы и проводятся эксперименты с целью создания открытых систем. Он обходит традиционную аргументацию «сверху вниз» и «снизу вверх», используя оба механизма. Процедуры «снизу вверх» используются для создания пространств возможностей, а методы «сверху вниз» отбирают структуры, которые функционируют хуже всего. Нисходящими механизмами могут быть как окружающая среда, так и люди, управляющие процессами развития.

Ключевые слова: генеративный искусственный интеллект, поддержка принятия решений, создание контента, информационные системы.

ANNOTATION

Closed systems of modern artificial intelligence, it seems, will not lead to the creation of intelligent machines in the near future. What is needed is open systems with nonlinear properties to create interesting properties for the artificial intelligence framework. This is the field of generative artificial intelligence, in which mechanisms are implemented and experiments are conducted in order to create open systems. It

bypasses the traditional "top-down" and "bottom-up" argumentation by using both mechanisms. Bottom-up procedures are used to create spaces of opportunity, and top-down methods select the structures that function the worst. The top-down mechanisms can be both the environment and the people who control the development processes.

Keywords: *generative artificial intelligence, decision support, content creation, information systems*

ВВЕДЕНИЕ

Инновация - это взять две существующие вещи и соединить их по-новому. В течение долгого времени в истории преобладало мнение, что художественные, творческие задачи, такие как написание стихов, создание программного обеспечения, дизайн одежды, сочинение песен, могут быть выполнены только человеком. С развитием искусственного интеллекта (ИИ), способного генерировать новый контент таким образом, что его уже невозможно отличить от человеческого мастерства, это представление кардинально изменилось.

Термин "генеративный ИИ" относится к вычислительным технологиям, способным генерировать на основе обучающих данных новый, на первый взгляд, осмысленный контент, такой как текст, изображения или аудио. Широкое распространение этой технологии на примере таких систем, как Dall-E 2, GPT-4 и Copilot, в настоящее время революционизирует наши методы работы и общения друг с другом. Системы генеративного ИИ могут использоваться не только в художественных целях для создания новых текстов, подражая писателям, или новых изображений, подражая иллюстраторам, но они могут и будут помогать человеку в качестве интеллектуальных систем, отвечающих на вопросы.

МЕТОДОЛОГИЯ

Под генеративной моделью ИИ понимается генеративное моделирование, которое сформировано архитектурой машинного обучения, следовательно, может создавать новые образцы данных на основе изученных закономерностей. Кроме

того, система генеративного ИИ включает в себя всю инфраструктуру, в том числе модель, обработку данных и компоненты пользовательского интерфейса. Модель выступает в качестве основного компонента системы, который облегчает взаимодействие и применение в более широком контексте.

Генеративная модель ИИ - это разновидность раскрытого машинного обучения, которая использует алгоритмы ИИ для создания новых экземпляров данных, опираясь на закономерности и взаимосвязи, наблюдаемые в обучающих данных. Генеративная модель ИИ имеет критически важный, но неполный характер, поскольку требует дальнейшей доработки под конкретные задачи с помощью систем и приложений.

Глубокие нейронные сети особенно хорошо подходят для целей генерации данных, тем более что глубокие нейронные сети могут быть разработаны с использованием различных архитектур для моделирования различных типов данных, например, последовательных данных, таких как человеческий язык, или пространственных данных, таких как изображения.

Большие генеративные модели ИИ, способные всесторонне и многогранно моделировать вывод в конкретных доменах или конкретных типах данных, часто называют также фундаментными моделями. В силу своего размера они обладают двумя ключевыми свойствами: эмерджентностью, то есть поведение часто неявно индуцируется, а не явно конструируется (например, модели GPT могут создавать записи календаря в формате, хотя такие модели не были явно обучены этому), и гомогенизацией, когда широкий спектр систем и приложений может теперь работать на основе единой, консолидированной модели (например, Copilot может генерировать исходный код на широком спектре языков программирования).

Можно использовать обе модальности для генерации изображений. В связи с этим генеративные модели ИИ также можно разделить на унимодальные и мультимодальные. Унимодальные модели получают инструкции от того же типа входа, что и их выход (например, текст). С другой стороны,

мультимодальные модели могут получать входные данные из различных источников и генерировать выходные данные в различных формах. Мультимодальные модели существуют для различных модальностей данных, например, для текста, изображений и аудио. Яркими примерами являются Stable Diffusion для генерации текста в изображение, MusicLM для генерации текста в музыку, Codex и AlphaCode для генерации текста в код, и, как уже упоминалось выше, GPT-4 для генерации изображения в текст, а также текста в текст.

Процедуры обучения, лежащие в основе различных генеративных моделей ИИ, существенно различаются. Например, генеративные состязательные сети (GAN) обучаются с помощью двух конкурирующих задач, одна из которых заключается в создании новых синтетических образцов, а другая - в выявлении синтетических образцов из реальных обучающих образцов таким образом, чтобы распределение синтетических образцов в конечном итоге было близко к распределению обучающих образцов. В других системах, таких как ChatGPT, основанных на разговорных моделях, используется обучение с подкреплением на основе обратной связи с человеком (RLHF). RLHF, используемый в ChatGPT, состоит из трех этапов: сначала создаются демонстрационные данные для подсказок, затем пользователи оценивают качество различных результатов для подсказки и, наконец, выучивается политика, которая генерирует желаемый результат с помощью обучения с подкреплением, чтобы этот результат получил хорошие оценки при ранжировании.

ОБСУЖДЕНИЯ

Далее мы рассмотрим четыре основные границы генеративного ИИ, которые, по нашему мнению, являются важными ограничениями в реальных приложениях. Следующие ограничения носят технический характер, поскольку относятся к тому, как текущие модели генеративного ИИ делают выводы, и, следовательно, ограничения возникают на уровне модели. В связи с этим, вероятно, ограничения сохранятся в долгосрочной перспективе и будут иметь последствия на уровне системы и приложений.

Кроме того, результаты работы генеративного ИИ, особенно LLM, как правило, нелегко поддаются проверке. Корректность моделей генеративного ИИ сильно зависит от качества обучающих данных и соответствующего процесса обучения. Системы и приложения генеративного ИИ могут реализовывать проверку корректности для предотвращения определенных результатов. Однако из-за "черного ящика" современных моделей ИИ использование таких систем в значительной степени зависит от доверия пользователей к достоверным результатам. Закрытый исходный код коммерческих готовых систем генеративного ИИ усугубляет этот факт и не позволяет проводить дальнейшую настройку и переобучение моделей. Одним из решений, позволяющих устранить последствия некорректных результатов, является использование генеративного ИИ для создания пояснений или ссылок, которые затем могут быть проверены пользователями. Однако такие объяснения опять же носят вероятностный характер и, следовательно, подвержены ошибкам; тем не менее, они могут помочь пользователям в принятии решений о том, когда следует принимать результаты работы генеративного ИИ, а когда нет.

Нарушение авторских прав. Генеративные модели, системы и приложения ИИ могут привести к нарушению законов об авторском праве, поскольку они могут создавать результаты, напоминающие или даже копирующие существующие произведения без разрешения или компенсации оригинальным авторам. Здесь можно выделить два потенциальных риска нарушения авторских прав. С одной стороны, генеративный ИИ может создавать нелегальные копии произведений, нарушая тем самым право авторов на воспроизведение. В частности, это может произойти, когда генеративный ИИ обучался на оригинальном контенте, защищенном авторским правом, но при этом он создает копии. Отсюда следует, что обучающие данные для создания систем генеративного ИИ должны быть свободны от авторских прав. Важно отметить, что нарушение авторских прав может происходить даже тогда, когда генеративный ИИ никогда не видел произведения, защищенного авторским

правом. С другой стороны, это открывает широкие возможности для исследований, особенно для исследователей BISE (Business & Information Systems Engineering) благодаря их междисциплинарному опыту.

Несмотря на огромный прогресс, достигнутый в последние годы, ряд аналитических и технических вопросов, связанных с развитием генеративного ИИ, еще не решен. Один из открытых вопросов связан с тем, как генеративный ИИ может быть эффективно адаптирован к конкретным областям применения и таким образом повысить производительность за счет более высокой степени определения контекста. Например, необходимы новые и масштабируемые методы настройки разговорных агентов на основе генеративного ИИ для применения в медицине или финансах. Это будет иметь решающее значение на практике для решения конкретных задач, связанных с BISE, где настройка может дать дополнительный прирост производительности. Новые методы настройки должны быть разработаны таким образом, чтобы обеспечить безопасность служебных данных и предотвратить их разглашение. Кроме того, для оперативного инжиниринга необходимы новые структуры, разработанные с учетом интересов пользователя и, следовательно, способствующие интерпретируемости и удобству использования.

Генеративный ИИ может автоматизировать различные задачи в области маркетинга и СМИ, где требуется создание контента, которые могут быть изучены в будущих исследованиях. Кроме того, генеративный ИИ может быть использован в рекомендательных системах для повышения эффективности распространения информации за счет персонализации, так как контент может быть лучше адаптирован к возможностям получателя.

Помимо разработки реальных систем и приложений, сообщество BISE должно также изучить возможности использования генеративного ИИ для поддержки творческих задач при создании новых дизайн-проектов. В этой связи перспективным направлением может стать включение генеративного ИИ в дизайн-мышление и аналогичные методологии, позволяющие объединить

творческие способности человека с вычислительными способностями. Это может способствовать поддержке различных фаз и этапов инновационных проектов, таких как генерация идей, выявление потребностей пользователей, создание прототипов, оценка дизайна и автоматизация проектирования, в которых различные типы моделей и систем генеративного ИИ могут использоваться и комбинироваться друг с другом для создания приложений для творческих индустрий. Если генеративный ИИ используется для совместного создания инновационных результатов, он также может обеспечить более эффективное повторное изучение различных видов проектной деятельности для обеспечения необходимого обучения.

ЗАКЛЮЧЕНИЕ

Генеративный ИИ - это направление ИИ, способное создавать новый контент, такой как тексты, изображения или аудио, который все чаще уже невозможно отличить от человеческого мастерства.

По этой причине генеративный ИИ способен изменить сферы и отрасли, которые зависят от творчества, инноваций и обработки знаний. В частности, он позволяет создавать новые приложения, которые ранее были невозможны или нецелесообразны для автоматизации, такие как реалистичные виртуальные помощники, персонализированное образование и обслуживание, цифровое искусство. Таким образом, генеративный ИИ имеет существенное значение для практиков и ученых BISE как междисциплинарного исследовательского сообщества. В нашей статье мы предложили концептуализацию принципов генеративного ИИ на уровне моделей, систем и приложений, а также на социально-техническом уровне и описали ограничения современного генеративного ИИ и мы подчеркнули многообразие возможностей, которые предоставляет генеративный ИИ через призму дисциплины BISE.

ЛИТЕРАТУРА

1. Li Y, Choi D, Chung J, Kushman N, Schrittwieser J, Leblond R, Eccles T, Keeling J, Gimeno F, Dal Lago A et al (2022) Competition-level code generation with alphacode. *Science* 378(6624):1092–1097
2. Liu P, Yuan W, Fu J, Jiang Z, Hayashi H, Neubig G (2023) Pre-train, prompt, and predict: a systematic survey of prompting methods in natural language processing. *ACM Comput Surv* 55(9):1–35
3. Мухамадиева Кибриё Баходировна. (2023). ОБЗОР СОВРЕМЕННЫХ ТЕХНОЛОГИЙ ИСКУСТВЕННОГО ИНТЕЛЛЕКТА В ТВОРЧЕСКОЙ ИНДУСТРИИ. <https://doi.org/10.5281/zenodo.8023673>
4. Мухамадиева Кибриё Баходировна. Анализ исследований по применению искусственного интеллекта в высшем образовании / К. Б. Мухамадиева // Образование и проблемы развития общества. – 2020. – № 2(11). – С. 119-124. – EDN NDWAUM.
5. Мухамадиева К. Б. Искусственный интеллект в развитии молодежи // Образование и проблемы развития общества. – 2021. – №. 2 (15). – С. 27-33.