



ТРАНСФОРМАЦІЙНА РОЛЬ ШТУЧНОГО ІНТЕЛЕКТУ В ОСВІТІ

Занько Н.В., доцент, кафедра МВГС, УАД
Глуховецький П.А., магістр, Європейський університет

Інтеграція нейронних мереж у сферу освіти відкриває безліч можливостей для модернізації навчального процесу та покращення якості навчання. Водночас формуються справжні виклики для усієї системи освіти, які пов'язані з фальсифікацією фактів генеративним штучним інтелектом та з порушенням правил академічної доброчесності серед учасників освітнього процесу.

Досліджено деякі можливості та обмеження використання мовної моделі GPT при формуванні текстового матеріалу. Було сформовано велику кількість запитів, які пропонували ChatGPT-3.5 [1] та ChatGPT-4 [2] згенерувати тексти різної тематики, щоб дослідити явище «галюцинування», коли модель створює текст або інформацію, яка виглядає точною та достовірною, але не має основи в реальних фактах.

За результатами експерименту, можна відзначити, що галюцинації проявлялися у трьох формах: поширення недостовірної інформації, перекручування фактів та фабрикування неіснуючих понять. Модель генерувала правдоподібний текст, який, проте, був не завжди достовірний. Наприклад, моделі створювали фіктивні наукові факти або посилалися на вигадані джерела. Найчастіше явище «галюцинування» виникало в тих випадках, коли запит був сформульований некоректно чи надто узагальнено. Потенційна небезпека полягає в тому, що галюцинації від ChatGPT часто виглядають дуже переконливими. Система вміє добре імітувати структуру і стиль правдивого тексту, що вводить користувачів в оману.

Потрібно розуміти, що мовна модель не копіює частини тексту з інформаційних джерел, які використовувались для її навчання. Вона самостійно генерує інформацію, використовуючи семантичні відносини між словами в реченні для створення тексту, який би за стилістикою був подібний до написаного людиною.

Найбільш деструктивною формою галюцинацій виглядає створення неіснуючих понять. Модель продукує імена, фрази або ідеї, які звучать правдоподібно, але не узгоджуються з жодним існуючим фактом або прийнятим знанням. Наприклад, нейромережа може використовувати сфабриковане дослідження або неіснуючу технологію.

В той же час ChatGPT-3.5 та ChatGPT-4 були надзвичайно ефективні при допомозі розробки освітньої програми, плану лекції, курсу навчальної дисципліни, тестових матеріалів. Системи штучного інтелекту можуть значно полегшити процес створення освітнього контенту: генерувати зображення, відео або аудіо матеріали, розробляти структуру презентації на основі введених даних або вказаних параметрів.

Обговорюючи використання генеративного штучного інтелекту, важливо торкнутись питання можливості дотримання академічної доброчесності. На



сьогодні немає однозначного вирішення проблеми захисту авторських прав у контексті контенту, створеного штучним інтелектом, наприклад, згенерованого ChatGPT. Через відсутність явних посилань на джерела або авторів у своїх результатах, такі моделі штучного інтелекту можуть ненавмисно генерувати текстову інформацію, що копіює матеріали, захищені авторським правом, без належного зазначення авторства. Це становить значну проблему, особливо в академічному та творчому секторах, де збереження авторських прав є надзвичайно важливим для визнання прав інтелектуальної власності справжніх авторів.

Наприклад, коли користувач створює запит на огляд певної наукової теорії, модель може створити текст, який має разючу схожість із захищеним авторським правом академічним матеріалом, який є у вільному доступі, але при цьому GPT не вказує його походження. Проблеми з порушенням авторських прав можуть виникнути, якщо текст використовується без належного цитування та атрибуції. Такі дії є порушення норм авторського права та етичних стандартів.

Студенти активно залучають ChatGPT для виконання індивідуальних завдань, написання рефератів, формування літературних оглядів для дипломних робіт, підготовки відповідей під час виконання модулів і т.п. Розпізнавання тексту створеного штучним інтелектом є складним завданням. Уже розробляються конкретні інструменти для визначення авторства штучного інтелекту. Наприклад, GLTR, також відомий як Giant Language Model Test Room [3] чи детектори на основі RoBERTa [4].

З огляду на вищесказане можна стверджувати, що потрібно змінювати підхід до оцінювання знань здобувачів освіти. Робити акцент на вміння студента аргументувати свою відповідь. Ставити задачі порівняння та співставлення фактів, аналізу інформації, аналітики. В письмових завданнях уникати відкритих запитань, відповіді на які легко згенерує мовна модель. Натомість, використовувати тестові завдання для виявлення рівня знань по дисципліні.

Як показує практика, з одної сторони штучний інтелект допомагає виконувати багато дій, демонструючи високу продуктивність. З іншої – породжує ряд проблем у зв'язку з фальсифікацією даних та порушенням академічної доброчесності. На часі вибудовувати оптимальну стратегію взаємодії з генеративними моделями, розробити протоколи та правила, які допоможуть використовувати штучний інтелект з користю для навчання.

Список літератури

1. GPT-3.5. <https://platform.openai.com/docs/models/gpt-3-5-turbo>.
2. GPT-4. <https://openai.com/index/gpt-4/>.
3. Gehrmann Sebastian, Strobelt Hendrik, Rush Alexander M. Gltr. (2019). Statistical detection and visualization of generated text. arXiv preprint arXiv:1906.04043.
4. Roberta Liu Yinhan та ін. (2019). Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692.