

6. GitHub - hashtopolis/server: Hashtopolis - A Hashcat wrapper for distributed hashcracking [Електронний ресурс]. URL: <https://github.com/s3inlc/hashtopolis> (дата звернення: 18.04.2023).

7. GitHub - nesfit/fitcrack: A hashcat-based distributed password cracking system [Електронний ресурс]. URL: <https://github.com/nesfit/fitcrack> (дата звернення: 18.04.2023).

8. GitHub - jmmcatee/cracklord: Queue and resource system for cracking passwords [Електронний ресурс]. URL: <https://github.com/jmmcatee/cracklord> (дата звернення: 18.04.2023).

9. GitHub - mandiant/gocrack: GoCrack is a management frontend for password cracking tools written in Go [Електронний ресурс]. URL: <https://github.com/fireeye/gocrack> (дата звернення: 18.04.2023).

УДК 044.56.53(043.2)

ШТУЧНИЙ ІНТЕЛЕКТ СНАТ GPT: МОЖЛИВОСТІ ТА ЗАГРОЗИ ТЕХНОЛОГІЧНОЇ НОВИНКИ

Галич Євгенія¹, Павленко Владислав²

Національний авіаційний університет

¹ 7405781@stud.nau.edu.ua, ² 7328430@stud.nau.edu.ua

18 лютого 2023 року найгарячіша технологічна новинка світу – чат-бот на основі штучного інтелекту ChatGPT (Generative Pre-trained Transformer) – стала доступною для реєстрації в Україні. Минуло всього декілька місяців відтоді, як ChatGPT з'явився у відкритому доступі, а хвиля новин навколо нього не вщухає, підвищуючи його актуальність серед користувачів.

Метою даної роботи є розбір новітньої технології – чату GPT, аналіз його принципу роботи та ризиків використання.

Додаток, розроблений компанією Open AI – це роботи, призначені для спілкування. Вони автоматично відповідають на письмові запитання, часто у подібній до людської манері. Однак за шоком від можливостей машин, що замінюють людей у різних формах дискусії ховається багато загроз, які будуть описані нижче.

ChatGPT швидко набув шаленої популярності: 1 млн користувачів за п'ять днів, в той час, як соціальна мережа Instagram йшла до таких показників охоплення 2,5 місяці, Facebook – 10 місяців, Netflix – 3,5 роки. GPT Chat працює з великим обсягом наданих даних. Платформа використовує сотні гігабайт тексту, які вона обробляє в режимі реального часу для створення нового контенту. В результаті GPT Chat може імітувати стиль і тон раніше введенного контенту, дозволяючи створювати зв'язні та логічні речення. Можливості використання штучного інтелекту досить масштабні. Наприклад, одним із основних вмінь технологічної

новинки є обробка природної мови. Використання методів машинного навчання для розпізнавання та аналізу текстових даних; виявлення теми, відношення між словами та іншими мовними структурами є фішкою чату GPT. Крім того, машинне навчання можна використовувати для вирішення різних завдань, таких, як прогнозування часових рядів, класифікація зображень і кластеризація даних. Для цього додаток використовує штучні нейронні мережі, щоб навчитися розпізнавати та аналізувати вхідні дані. Ще одне використання GPT Chat – автоматична генерація відповідей на запитання користувачів. Відповіді можуть стосуватися різних сфер, таких, як здоров'я, фінанси, мода, кулінарія тощо.

Головною перевагою GPT Chat є його ефективність і швидкість. Платформа може обробляти великі обсяги даних і виконувати завдання швидше, ніж людина, та працювати без втоми і помилок, що характеризує роботу застосунку як надійну і точну. Використовуючи штучний інтелект, компанії та інші організації можуть скоротити витрати на робочу силу, оскільки він виконує багато рутинних і повторюваних завдань, які раніше виконувала людина.

Незважаючи на те, що ChatGPT є найбільш "просунутою" технологією з використанням штучного інтелекту на сьогоднішній день, вона далека від досконалості.

Застосунок також має свої недоліки та обмеження. Наприклад, обмеженість у прийнятті рішень та недостатній контроль. Хоча чат може дуже ефективно виконувати рутинні та повторювані завдання, все ж таки існує обмежена здатність при вирішенні складних проблем, які вимагають значного розуміння контексту та досвіду. Також, у роботі штучного інтелекту, може бути складно контролювати та розуміти його дії та рішення. Це може становити проблему у випадку, коли рішення мають серйозні наслідки.

Як і будь-який технологічний продукт, використання штучного інтелекту не є повністю безпечним. Наведемо декілька ризиків, пов'язаних з використанням ChatGPT:

1. *Поширення дезінформації*: GPT-розмови можуть давати неправдиві або оманливі відповіді, які можуть бути неправильно інтерпретовані або використані для поширення дезінформації.

2. *Підтримка упередженості*: оскільки модель навчається на великій кількості даних, навчальні дані можуть містити поширені упередження та патерни дискримінації. Якщо ці шаблони не виявити і не виправити, Chat GPT може увічнити негативне ставлення до певних груп.

3. *Вразливість до хакерських атак*: Chat GPT може бути вразливим до різних видів хакерських атак, зокрема атак типу "відмова в обслуговуванні" (DoS), спроб злому та крадіжки даних.

4. *Загроза конфіденційності*: GPT може зберігати персональні дані користувачів. Це може призвести до порушення конфіденційності або використання цих даних сторонніми особами чи організаціями.

Перспективи використання ChatGPT є дуже високими. Технології штучного інтелекту постійно розвиваються, що дозволяє застосунку вдосконалюватись та надавати більше корисної інформації та рішень, розглядатись у контексті спрощення і пришвидшувати деякі процеси нашого життя. ChatGPT, безсумнівно, стане в нагоді маркетологам при написанні рекламних текстів, інфлюенсерам при підготовці контенту для соціальних мереж і власникам бізнесу при веденні необхідної документації. Чат-боти значно спрощують процес пошуку інформації, надаючи користувачам розгорнуті відповіді замість переліку посилань. У майбутньому ChatGPT та інші розробки на основі штучного інтелекту можуть замінити людину у виконанні багатьох завдань. Однак, для того, щоб сформулювати правильні вимоги, вибрати найбільш задовільне рішення та оцінити, чи відповідають результати певним очікуванням, потрібні люди з високим рівнем експертизи та глибоким розумінням конкретної сфери спеціалізації. Також застосунок може допомогти у покращенні освіти, надавати корисні ресурси для навчання та підтримувати сам процес для студентів та вчителів, що є особливо важливим в будь-які часи. Загалом, існує великий потенціал для використання штучного інтелекту, його розвиток може значно покращити якість нашого життя та забезпечити розвиток різних галузей науки і технологій. Однак, важливо, щоб штучний інтелект використовувався етично і відповідально для надання безпеки, конфіденційності і довіри його користувачам.

Перевіримо штучний інтелект у дії. За допомогою сервісу було змодельовано принцип роботи атак на конфіденційні дані користувачів методом SQL ін'єкцій. Для цього було зроблено запит до чату GPT - "Принцип роботи атак на конфіденційні дані користувачів методом SQL ін'єкцій". Отримали наступне.

Розглянемо більш детально, як саме працює атака методом SQL-ін'єкцій. Зловмисник може використовувати SQL-ін'єкції для виконання інших зловмисних дій, таких як вставка, оновлення або видалення даних з бази даних. Наприклад, він може ввести наступну послідовність символів у поле логіну мовою CSS: `' ; DROP TABLE users;--` . Цей запит викличе помилку в базі даних та видалить всі дані з таблиці "users". Щоб запобігти атакам методом SQL-ін'єкцій, необхідно правильно валідувати вхідні дані та коректно оброблювати запити до бази даних. Крім того, рекомендується використовувати параметризовані запити, які дозволяють окремо вказувати дані та запит до бази даних, що допоможе уникнути можливих атак методом SQL-ін'єкцій.

В ході дослідження прийшли до висновку, що користувач, не маючи кваліфікацію, за допомогою ChatGPT може отримати доступ до будь-якого ресурсу шляхом експлуатації певних уразливостей в системі.

Науковий керівник – к.т.н., доцент Хохлачова Ю.Є.