УДК 004.8+004.49

ШТУЧНИЙ ІНТЕЛЕКТ ЯК ІНСТРУМЕНТ КІБЕРАТАК І КІБЕРБЕЗПЕКИ

ARTIFICIAL INTELLIGENCE AS A CYBERATTACK TOOL AND CYBERSECURITY

Юрій Каун[1], Оксана Собчук[2]

[1]*Луцький інститут розвитку людини Університету "Україна",*
вул. Георгія Гонгадзе, 5, м. Луцьк, 43027, Україна
[2]*Волинський національний університет імені Лесі Українки,*
просп. Волі, 13, Луцьк, 43025, Україна

Abstract. *The report discusses the possibilities of using artificial intelligence tools to create fakes, malware, phishing emails, and possible ways to combat them.*

Поява GPT та інших мовних моделей стала важливим етапом на шляху впровадження штучного інтелекту (ШІ) у життя сучасної людини. ШІ може допомогти користувачам з: швидким та релевантним пошуком інформації, творчими завданнями, персоналізацією навчання, аналізом великих обсягів даних, виявленням закономірностей, прискоренням темпів наукових відкриттів та розробки нових технологій. ШІ стрімко змінює світ розробки програмного забезпечення, шляхом автоматизації рутинних і повторюваних завдань. Деякі інструменти ШІ здатні генерувати к1од на основі простих описів або природної мови. Це може значно прискорити процес розробки та зробити його доступнішим для людей з меншим досвідом програмування.

Попри те, що моделі ШІ мають великий потенціал для позитивного застосування, зловмисники також почали використовувати їх для вдосконалення своїх кіберзлочинних методів. Недавнє дослідження, проведене вченими Університету Індіани, виявило понад 200 вкрадених і зламаних відкритих мовних моделей, що пропонуються для хакерської діяльності. Це свідчать про зростаючу тенденцію використання ШІ для зловмисних цілей. Зокрема, створення дипфейків на основі цілком реалістичних та оманливих відео, які можуть мати значний вплив на громадську думку та хід подій. Зневажливі відео, підроблені документи та фейкові акаунти в соціальних мережах можуть бути використані для маніпулювання громадськістю, підриву довіри до інституцій та розпалювання соціальних заворушень.

ШІ може генерувати дуже переконливі фішингові електронні листи, які важко відрізнити від легітимних повідомлень. Він усуває орфографічні, граматичні та друкарські помилки, які часто зустрічаються у фішингових електронних листах, створених людьми. Це робить їх ще більш реалістичними та оманливими. Виявлення шкідливого ПЗ та електронних листів, створених за допомогою штучного інтелекту, стає все складнішим завданням. Це пов'язано з тим, що ШІ постійно розвивається, а зловмисники все частіше використовують його для створення більш складних та оманливих загроз.

Більшість відомих ШІ має етичні обмеження, що ускладнюють їх шкідливе використання. Проте, їх можна обійти за допомогою уникнення прямих запитів, розбивці завдань, використання нейтральної мови. Тому зловмисники можуть успішно використовувати ШІ для створення шкідливого коду, не порушуючи явних етичних обмежень моделі. При цьому інструменти ШІ можуть використовуватися людьми з мінімальними технічними знаннями, що полегшує діяльність кіберзлочинців, оскільки їм не потрібні глибокі знання програмування чи доступ до дорогих ресурсів.

Очікується, що використання ШІ призведе до зростання кількості та складності кібератак, таких як атаки програм-вимагачів, фішинг та розповсюдження шкідливих програм. Організаціям буде складніше захиститися від кібератак, створених за допомогою ШІ, що може призвести до значних фінансових втрат, включаючи викупні платежі, витрати на

відновлення та втрату репутації. Кібератаки можуть призвести до порушення роботи систем та послуг, що негативно впливатиме на клієнтів та партнерів.

Використання генеративного ШІ для створення зловмисного ПЗ та фішингових електронних листів стає серйозною загрозою кібербезпеці. Ці інструменти, навчені обходити методи виявлення, ставлять перед фахівцями з кібербезпеки нові виклики. Адже зловмисники навчають моделі ШІ алгоритмам виявлення шкідливого програмного ПЗ. Модель ШІ використовує знання про методи виявлення, щоб створювати шкідливе ПЗ, яке уникає цих методів.

Хоча прямого доказу того, що шкідливе ПЗ чи електронний лист було створено за допомогою ШІ, може бути й недостатньо, існують методи, які дозволяють виявити такі загрози. Один із таких методів ґрунтується на використанні інструментів ШІ для сканування тексту, який підозрюється у створенні за допомогою ШІ.

Фахівці з кібербезпеки можуть використовувати ШІ, щоб імітувати методи хакерів, ставлячи подібні запитання та аналізуючи відповіді. Цей процес може допомогти їм виявити потенційні вразливості в мережі, про які вони могли не знати та розробити превентивні заходи та зміцнити захист мережі.

Зростаюча загроза дезінформації та пропаганди у кіберпросторі потребує комплексної стратегії протидії, яка має ґрунтуватися на ретельному моніторингу, проактивних попередженнях і активній співпраці. Важливо використовувати передові методи моніторингу на основі ШІ та машинного навчання для раннього виявлення потенційних кампаній дезінформації, аналізуючи великі обсяги даних із соціальних мереж, веб-сайтів, новинних ресурсів та форумів. Це дозволить виявляти аномалії та підозрілі моделі поведінки, ідентифікуючи ботів, тролів та інших агентів дезінформації.

Також необхідно використовувати технічні та правові інструменти для блокування поширення дезінформаційного контенту та його видалення з онлайн-платформ, а також інформувати користувачів про надійність контенту і позначати фейкові новини. Важливо сприяти поширенню достовірної інформації з надійних джерел.

Ще один ключовий аспект – підвищення обізнаності та стійкості громадськості шляхом впровадження освітніх програм, які розвивають критичне мислення та вміння розпізнавати фейки; підтримка незалежних ЗМІ, які надають неупереджену інформацію, та залучення громадських організацій та активістів до боротьби з дезінформацією.

Світ дійсно вступає в епоху, де штучний інтелект використовується не лише для захисту, але й для атак. Фахівцям з кібербезпеки потрібно постійно оновлювати свої методи та інструменти, щоб випереджати зловмисників, які також використовують ШІ.

Бібліографія

1. Christopher Mouton. ChatGPT is creating new risks for national security. URL: https://www.defensenews.com/opinion/2023/07/20/chatgpt-is-creating-new-risks-for-national-security/ (дата звернення: 1.06.2024)

2. Zachary Folk, CEH, CISSP-ISSEP, Security+. What Enterprises Need to Know About ChatGPT and Cybersecurity/ URL: https://www.isaca.org/resources/news-and-trends/industry-news/2023/what-enterprises-need-to-know-about-chatgpt-and-cybersecurity (дата звернення: 1.06.2024)

3. Nas Ali. ChatGPT and how AI is changing cyber security. URL: https://thesecuritycompany.com/the-insider/chatgpt-and-how-ai-is-changing-cyber-security-2/ (дата звернення: 1.06.2024)