УДК 004.89

РОЗВИТОК ШТУЧНОГО ІНТЕЛЕКТУ ТА МАШИННОГО НАВЧАННЯ ОБ'ЄКТІВ ГОСПОДАРЮВАННЯ

Курепін В.М., к.е.н., доц.
Дідняк А.В., здобувач РВО бакалавр
Миколаївський національний аграрний університет
м. Миколаїв, Україна, didniak2108@cioud.com

Анотація. Вплив штучного інтелекту та машинного навчання на життя людей, як у побуті, так і на їхню роботу не визиває сумнівів. З часом, рік від ріку, зростає споживання продуктів та послуг, створених на основі штучного інтелекту та машинного навчання. Щоб захистити дані не тільки ваших клієнтів, а і самі данні й сам штучний інтелект та алгоритми від зловживання від тролінгу та порушення працездатності необхідно на постійній основі вживати спеціальні заходи.
Ключові слова: зловмисники, ІТ-галузь, шкідливі атаки, критичне розуміння, штучний інтелект, ефективність

Досліджуючи тему нашої доповіді, ми дійшли висновку, що існують практичні проблеми у забезпеченні безпеки при розробці продуктів та використанні веб-служб, заснованих на штучному інтелекті, які необхідно вирішити зараз. Наголошуємо, превентивні заходи, які вирішуються на об'єктах господарювання в рамках корпоративної ІТ-галузі (забезпечення безпеки клієнтів та захист їх даних) є стратегічною проблемою у довгостроковій перспективі [1, с. 68].

Ні для кого немає сумнівів, що зловмисниками постійно використовуються штучний інтелект, на його основі здійснюються високоскладні, творчі та шкідливі атак, будь то окремі тролі чи цілі вовчі пакети. У зв'язку з цим виникають проблеми захисту продуктів та служб на основі штучного інтелекту.

Ми позначили проблеми, що мають стратегічне значення для галузевих фахівців, які керують розробкою систем безпеки об'єктів господарювання. Розглянемо результати нашого дослідження, які показують наступне:

а) для усунення проблем безпеки потрібна адаптація існуючих практик безпеки з урахуванням особливостей штучного інтелекту та машинного навчання;

б) деякі моделі машинного навчання здебільшого не можуть розрізняти шкідливу вхідну інформацію та нешкідливі нестандартні дані. Зловмисникам не потрібно компрометувати набори даних, якщо вони є вільними для участі в них [2, с. 129]. Згодом шкідливі дані стають надійними, якщо структура та форматування даних залишається правильною;

в) враховуючи велику кількість прихованих класифікаторів, які можна використовувати в моделі глибокого навчання, надто багато довіри ставиться на вихідні дані процесів та алгоритмів прийняття рішень штучним інтелектом/машинним навчанням без критичного розуміння того, як ці рішення

були досягнуті. Таке спотворення унеможливлює алгоритмів штучного інтелекту та машинного навчання продемонструвати логіку своєї роботи і ускладнює доказ правильності результатів, коли вони ставляться під сумнів;

г) штучний інтелект та машинне навчання все частіше використовуються для підтримки процесів прийняття важливих рішень у виробничій безпеці, де помилка може призвести до серйозних травм чи летального випадку. Відсутність можливостей отримати аналітичну звітність про роботу алгоритмів штучного інтелекту та машинного навчання не дозволяє використовувати цінні дані як докази у вирішенні питань з розслідування нещасних випадків на виробництві.

З деякими з описаних у нашому дослідженні проблемами розробники IT-технологій вже впоралися/впораються протягом найближчих років [3, с. 212]. Якщо на цю область безпекових заходів не звертати увагу, ми ризикуємо майбутнім штучним інтелектом стати чорною скринькою через нашу нездатність довіряти чи розуміти (і змінювати за потреби) процеси прийняття рішень штучним інтелектом. З точки зору безпеки це фактично означає втрату контролю.

Метою та цілями нашого дослідження полягають у тому, щоб виділити питання інженерії безпеки, які є унікальними для області штучного інтелекту та машинного навчання, виявити деякі думки та спостереження щодо виникаючих загроз і поділитися напрацюваннями про потенційні виправні заходи.

Нові завдання під час проектування систем безпеки. Традиційні вектори атак програмного забезпечення є критично важливими для вирішення, але вони не забезпечують достатнього охоплення загроз штучним інтелектом та машинним навчанням. IT-індустрія має уникати вирішення проблем нового покоління за допомогою застарілих методів. Але, важливо створити нову інфраструктуру на основі нових підходів, які зможуть усунути недоліки у розробці та експлуатації служб на основі штучного інтелекту та машинного навчання. Зупинимося на них:

1. Основи розробки систем безпеки повинні включати в себе концепції стійкості та вибірковості. Приділяючи більше уваги питанням, пов'язаним з штучним інтелектом, необхідні інвестиції у цю галузь [4, с. 212].

2. Штучний інтелект повинен уміти розрізняти навмисні відхилення у поведінці інших, але при цьому не допускати впливу цих відхилень на власні механізми взаємодії з людьми. Для цього потрібне загальне і постійне розуміння забобонів, стереотипів, жаргону та інших культурних конструктів. Це допоможе захистити штучний інтелект від атак на набір даних [5, с. 68].

3. Алгоритми машинного навчання повинні бути здатні відрізняти шкідливо введені дані від нешкідливих подій, відхиляючи навчальні дані, які негативно впливають на результати. В іншому випадку моделі навчання завжди схильні до ігор зловмисниками і тролями.

4. Штучний інтелект повинен містити у собі вбудовані функції аналітичної експертизи. Це дозволяє підприємствам надавати клієнтам прозорість та підзвітність їх штучного інтелекту, гарантуючи, що його дії не тільки перевіряються, а й юридично виправдані.

5. Штучний інтелект повинен виявляти та захищати конфіденційну інформацію, навіть якщо люди не впізнають її такою [6, с. 31].

Отже, створений штучний інтелект приносить користь людству. Повсюдне використання штучного інтелекту породжує серйозні питання безпеки. Наскільки ми можемо бути впевнені у безперебійності та безперервності роботи критично важливих систем, керованих штучним інтелектом? Чи можна довірити машині здоров'я та життя людини? Однозначно, технології штучного інтелекту стрімко покращують життя людства – від повсякденних завдань до глобальних бізнес-процесів. Ці технології мають величезний потенціал для автоматизації рутини, підвищення ефективності, генерації ідей, народження інновацій та створення нових можливостей у безпекової галузі.

Список літератури

1. Курепін В. М. Іваненко В. С. Застосування цифрових технологій у сільському господарстві для досягнення цілей сталого розвитку. Modern Economics. 2024. № 47(2024). С. 62-69. DOI: https://doi.org/10.31521/modecon.V47(2024)-09.

2. Бацуровська І. В., Кашина Г. С., Курепін В. М. Інтеграція сучасних освітніх технологій, системи якості вищої освіти та принципів безпеки життєдіяльності у підготовці фахівців. *Перспективи та інновації науки. Серія : Педагогіка. Психологія. Медицина.* 2025. № 3(49). С. 119-136. URL: https://dspace.mnau.edu.ua/jspui/handle/123456789/20873.

3. Kurepin V. M., Oliynyk T.G. Use of artificial intelligence in tourist activities. Research and education in the global world: eurointegration processes during the wartime : Book of papers of the X International Forum for Young Researchers, Kharkiv, May 10, 2024 / O. M. Beketov National University of Urban Economy in Kharkiv, TESOL-Ukraine. Kharkiv : O. M. Beketov NUUE in Kharkiv, 2024. P. 211-212. URL: https://dspace.mnau.edu.ua/jspui/handle/123456789/17943.

4. Лотарєва Д. Використання інноваційних технологій та методів управління виробничими процесами за допомогою штучного інтелекту // Молодь, наука, бізнес : матеріали Всеукр. інтер.-конф. здоб.вищ.освіти і мол.учених, 5-6 жовтня 2022 р., м. Миколаїв. Миколаїв : МНАУ, 2022. С. 77-80. URL: https://dspace.mnau.edu.ua/jspui/handle/123456789/11860.

4. Іваненко В. С., Курепін В. М. Наближення національного законодавства до міжнародних норм з питань безпеки праці // OSHAgro – 2023: матеріали III Міжнародної науково-практичної конференції (м. Київ, 3 жовтня 2023 р.). Київ: НУБіП України, 2023. С. 66-69. URL: https://dspace.mnau.edu.ua/jspui/handle/123456789/15934.

5. Іваненко В. С. Оптимізація асортименту плодоовочевої продукції в умовах кризи за допомогою штучного інтелекту // Сучасні підходи до вирощування, переробки і зберігання плодоовочевої продукції : матеріали міжнародної науково-практичної конференції, 17 листопада 2022 р., м. Миколаїв. Миколаїв : МНАУ, 2022. С. 30-32. URL: https://dspace.mnau.edu.ua/jspui/handle/123456789/12135.