Штучний інтелект та кібербезпека

УДК 004.08

Владислав Орбан[1]

Західноукраїнський національний університет, [1]v.orban@st.wunu.edu.ua

Сьогодні ми є свідками того, як технології формують не лише бізнес-середовище, а й архітектуру глобальної безпеки. І немає більш визначальної сили в сучасному цифровому світі, ніж штучний інтелект. Ще зовсім недавно системи кіберзахисту базувалися на сигнатурних методах, жорстких правилах брандмауерів та ручному аналізі логів. Зараз же ми спостерігаємо перехід до аналізу великих даних за допомогою різних моделей. ШІ став одночасно і помічником і загрозою для захисту інформаційних систем, і наша ключова мета як фахівців — зрозуміти, як ефективно володіти цим інструментом, щоб захистити нашу інфраструктуру від загроз нового покоління.

Перевагами ШІ є: традиційні системи виявлення вторгнень можуть генерувати лавини хибних спрацьовувань, серед яких буває важко виявити справжній інцидент. Саме тут на допомогу приходять алгоритми машинного навчання, де замість постійного оновлення сигнатур вірусів, модель тренується на масивах даних нормальної роботи мережі, і може виявляти навіть найменші відхилення, такі як: нестандартне використання протоколів, або ж спроби тунелювання трафіку, це дозволяє фіксувати Zero-Day атаки, тобто ті, що не відомі, наприклад, антивірусу. Штучний інтелект здатний не лише виявляти загрозу, а й автономно реагувати на неї. У разі фіксації інциденту система може миттєво ізолювати скомпрометований вузол, динамічно переналаштувати правила VLAN для блокування сегмента мережі або ж зупинити процеси, що виглядають підозріло, до того як спеціаліст з кібербезпеки встигне відкрити сповіщення про небезпеку.

Сучасні рішення використовують глибинне навчання для аналізу поведінки програм у оперативній пам'яті комп'ютера, ефективно блокуючи програми-вимагачі (Ransomware) на етапі спроби масового шифрування файлів. Але не варто забувати, що ШІ технології можуть допомогти хакерам зі зламами систем [2]. Через невпинний розвиток технологій, ми стикаємося з масштабуванням та ускладненням кібератак [3]. Великі мовні моделі дозволяють автоматизувати написання фішингових повідомлень. Зловмиснику не потрібно вивчати жертву власноруч, адже ШІ агрегує дані з відкритих джерел і генерує листи, які відрізнити від легітимних комунікацій майже неможливо. Ну і не варто забувати про технологію Deepfake, яку вже зараз використовують для омани корпоративних співробітників з метою несанкціонованого переказу коштів. Згідно даних представлених на світовому економічному форумі [1] загроза для ШІ становить 87%.

На фоні зростання масового використання технологій штучного інтелекту як і в роботі бізнесу, так і в звичайному житті, з'явилися віруси, які можуть адаптуватися до будь-якого середовища. За допомогою методів машинного навчання, такий код розуміє, коли він знаходиться в ізольованому середовищі для аналізу і «засинає», активуючи свій деструктивний функціонал лише після потрапляння на реальну робочу станцію.

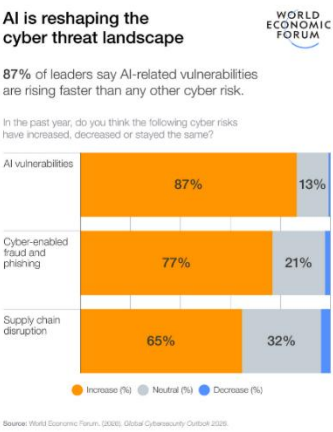

Рис.1. Дослідження про загрози для ШІ від World Economic Forum

За допомогою ШІ можна за лічені хвилини проаналізувати величезні масиви коду або конфігурацій, виявляючи всі приховані вразливості, такі як SQL-ін'єкції або ж помилки в налаштуваннях доступу, що прискорює етап розвідки в десятки разів.

У зв'язку з тим, що все більше систем покладаються на системи безпеки на базі ШІ, зловмисник може атакувати саме його. Найнебезпечнішим методом є «отруєння даних». Таким чином, якщо зловмисник отримує доступ до датасетів, на яких тренується модель, він може непомітно внести зміни так, щоб навчити модель сприймати шкідливий трафік за безпечний. Це означає, що відтепер інженерам з кібербезпеки необхідно захищати не лише сервери, бази даних та мережеве обладнання. Ми маємо захищати саму математику, забезпечуючи цілісність наборів даних та стійкість ML-моделей до маніпуляцій.

Підсумовуючи, хочу зазначити, що штучний інтелект навряд зможе повністю замінити експертів з кібербезпеки, однак ті спеціалісти, які володіють знаннями та інструментарієм Data Science, зможуть з легкістю замінити тих, хто відмовляється адаптовуватися до нових реалій. У недалекому майбутньому кіберпростір належатиме тим, хто зможе найшвидше аналізувати дані та адаптуватися до нестандартних викликів нашого часу.

1. World Economic Forum. (2025, January 14). Global cybersecurity outlook 2026: Infographics. https://www.weforum.org/publications/global-cybersecurity-outlook-2026/infographics-global-cybersecurity-outlook-2026/
2. Deepstrike. (2025, January 1). Cybercrime statistics 2025: Trends, impact, and insights. https://deepstrike.io/blog/cybercrime-statistics-2025
3. Programs.com. (2024, December 31). AI cyberattack statistics: The role of artificial intelligence in modern cyber threats. https://programs.com/resources/ai-cyberattack-stats/