DOI 10.15276/imms.v15.no1.115
УДК: 004.8:004.056

Informatics and Mathematical Methods in Simulation
Vol.15 (2025), No. 1, pp. 115-125

ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ В КІБЕРБЕЗПЕЦІ

Б.В. Приступа [1,2], Н.В. Герасимюк [2], Я.В. Рожковський[2]

[1]Військовий медичний клінічний центр Південного регіону
2, Пироговська вул., Одеса, 65044, Україна
[2]Одеський національний медичний університет
2, Валіховський провулок, Одеса, 65082, Україна
Email: bogdan.prystupa@onmedu.edu.ua

Сучасний розвиток цифрових технологій, таких як Інтернет речей, штучний інтелект та хмарні обчислення, сприяє зростанню обсягів даних, які потребують ефективного захисту. Штучний інтелект є перспективним напрямком у сфері кібербезпеки, оскільки дозволяє автоматизувати аналіз великих обсягів даних, ідентифікувати аномалії в трафіку та швидко реагувати на загрози. Основні результати дослідження включають визначення основних алгоритмів машинного навчання, що застосовуються у сфері кібербезпеки, серед яких підтримуючі векторні машини, дерева рішень, нейронні мережі, а також підходи, що базуються на підкріплювальному навчанні. Доведено, що використання глибокого навчання дозволяє досягати точності виявлення загроз до 96 відсотків, що перевищує традиційні методи аналізу кіберзагроз. Розглянуто проблеми, пов'язані з впровадженням штучного інтелекту у сфері кібербезпеки, зокрема вразливість моделей до атак на навчання, що можуть змінювати поведінку алгоритмів та обходити системи захисту. Оцінено регуляторні виклики та необхідність створення нормативних актів для контролю технологій автономних штучних інтелектуальних систем, що працюють поза контролем офіційних IT-структур організацій. Цінність дослідження полягає у розробці комплексного підходу до використання штучного інтелекту у сфері кібербезпеки, що дозволить підвищити рівень захисту інформаційних систем та мінімізувати ризики кібератак. Отримані результати сприятимуть розширенню наукових підходів у галузі інтелектуальних систем безпеки та розробці ефективних алгоритмів виявлення загроз у реальному часі. Практичне значення проведеного дослідження полягає у можливості застосування отриманих результатів для вдосконалення систем виявлення вторгнень, антивірусного програмного забезпечення, а також впровадження автоматизованих рішень для захисту критичних інфраструктур від кібератак.

Ключові слова: штучний інтелект, машинне навчання, кібербезпека виявлення загроз, аналіз аномалій, глибоке навчання.

Вступ. Стрімкий розвиток інтелектуальних технологій, Інтернету речей (IoT) та комп'ютерних пристроїв призвів до створення величезних обсягів даних, які потребують належного рівня захисту [1]. Хоча ці інновації значно спростили повсякденне життя та оптимізували бізнес-процеси, вони водночас спричинили нові виклики у сфері кібербезпеки, підвищивши ризик кібератак та витоку конфіденційної інформації [2].

Згідно зі звітом Cybersecurity Ventures (2023), кібервитрати у світі прогнозовано досягнуть 8 трильйонів доларів США у 2023 році, що робить кіберзлочинність третьою за величиною економікою після США та Китаю. Очікується, що глобальні витрати на кіберзлочинність зростатимуть на 15% щорічно протягом наступних трьох років, досягаючи 10,5 трильйонів доларів США щорічно до 2025 року [3]

Найпоширенішими типами атак залишаються фішинг, шкідливе програмне забезпечення, атаки на відмову в обслуговуванні (DoS), експлойти нульового дня та методи соціальної інженерії [4, 5].

Такі загрози несуть серйозні ризики для приватних осіб, підприємств і великих організацій, підриваючи їхню інформаційну безпеку та стабільність [6]. Останніми роками кібербезпека зазнає значних змін як у технологічних, так і в робочих аспектах

комп’ютерних систем [6]. Штучний інтелект є рушійною силою цих змін, оскільки машинне навчання (ML) та глибоке навчання (DL) є фундаментальними компонентами ШІ, які відіграють вирішальну роль у розкритті інформації з даних [7].

Штучний інтелект дозволяє аналізувати величезні обсяги даних, виявляти загрози в реальному часі та автоматизувати процеси захисту інформаційних систем. Це робить його невід'ємною частиною сучасних систем кібербезпеки. Проте, зростання кількості автономних інструментів ШІ створює новий феномен — Shadow AI. Це поняття означає використання несанкціонованих або нерегульованих систем штучного інтелекту, які працюють поза контролем офіційних IT-структур організацій. Shadow AI може сприяти підвищенню продуктивності, проте також створює значні ризики для безпеки даних та конфіденційності [8].

Таким чином, сучасні виклики у сфері кібербезпеки вимагають інтеграції передових рішень на основі штучного інтелекту, що дозволяють не лише виявляти загрози, а й забезпечувати адаптивний захист у реальному часі.

Мета дослідження. Метою цієї роботи є аналіз сучасних підходів до використання штучного інтелекту в кібербезпеці, виявлення ключових методів та технологій, що застосовуються для захисту інформаційних систем, а також оцінка їхньої ефективності. Зокрема, увага приділяється методам машинного навчання, аналізу поведінки, виявленню аномалій та системам автоматизованої відповіді на інциденти.

Для досягнення поставленої мети було використано методи системного аналізу, огляд наукової літератури, а також аналіз сучасних технологій штучного інтелекту, що застосовуються в кібербезпеці. Штучний інтелект у кібербезпеці відіграє ключову роль у забезпеченні захисту інформаційних систем. Його застосування дозволяє значно підвищити рівень виявлення загроз, скоротити час реагування на інциденти та знизити ризики фінансових втрат [9]. Системи на основі машинного навчання здатні аналізувати величезні обсяги даних, виявляючи потенційно небезпечну активність ще на етапі її зародження.

Інструменти та методи штучного інтелекту в кібербезпеці. За словами Камачо (2024), найпоширеніші методи AI, що використовуються в кібербезпеці, включають машинне навчання (Machine Learning) і глибоке навчання (Deep Learning) [10]. ML — це підгалузь штучного інтелекту, яка зосереджена на розробці алгоритмів і статистичних моделей, які дозволяють комп’ютерам вчитися та робити прогнози або рішення на основі [11]. Існує три основні типи ML, а саме: контрольоване навчання, неконтрольоване навчання та навчання з підкріпленням.

Глибоке навчання (DL) представляє собою напрям машинного навчання (ML), що використовує багатошарові штучні нейронні мережі (ANN) для виявлення складних взаємозв'язків та закономірностей у даних. Навчання відбувається на основі нейронної архітектури з кількох рівнів, яка включає вхідний шар, один або більше прихованих шарів та вихідний шар [11].

Застосування штучного інтелекту у сфері кібербезпеки значно підвищує ефективність виявлення загроз та реагування на інциденти. Завдяки алгоритмам машинного навчання можна аналізувати великі обсяги даних, розпізнавати аномалії та автоматизувати процеси захисту інформаційних систем. Різні методи ШІ використовуються для моніторингу мережевого трафіку, виявлення вторгнень, класифікації загроз та розпізнавання фішингових атак. Розглянемо основні інструменти та методи, що застосовуються для посилення безпеки в цифровому середовищі.

Системи виявлення вторгнень. У системах виявлення вторгнень (Intrusion detection systems англ) широко застосовуються різні алгоритми машинного навчання для класифікації мережевого трафіку. Наприклад, SVM і KNN використовуються для ідентифікації трафіку як нормального або шкідливого. SVM знаходить оптимальну гіперплощину для розділення класів даних, тоді як KNN класифікує об'єкти на основі найближчих сусідів [12, 13]. Наївний Байєс діє як ймовірнісний класифікатор, що

передбачає незалежність ознак і визначає ймовірність належності пакета до категорії «нормальний» або «шкідливий» [14, 15]. Дерева рішень (DT), своєю чергою, моделюють процес прийняття рішень шляхом рекурсивного поділу простору ознак [16, 17]. Для неконтрольованого навчання застосовується K-Means, що групує трафік у кластери, допомагаючи ідентифікувати аномалії та нові загрози [18, 19]. Штучні нейронні мережі (ANN) здатні навчатися як у контрольованому, так і в неконтрольованому режимах, ідентифікуючи складні шаблони вторгнень шляхом налаштування ваг та зв'язків між нейронами [20]. AdaBoost, у свою чергу, підвищує точність класифікації, зосереджуючись на складних для класифікації прикладах [21].

Мережеві системи виявлення вторгнень. У Мережеві системи виявлення вторгнень (Network intrusion detection system англ) алгоритми типу документа використовуються для класифікації трафіку за допомогою правил, сформованих на основі історичних даних [22, 23]. Random Forest (RF) підвищує точність шляхом об’єднання прогнозів кількох дерев рішень [**Ошибка! Закладка не определена.**]. Серед моделей глибокого навчання для NIDS виділяються Convolutional Neural Networks (CNN). Вони аналізують корисне навантаження пакетів, вивчаючи ієрархічні представлення функцій, що дозволяє виявляти складні шаблони атак [24, 25].

Розпізнавання зображень та CAPTCHA. Для розпізнавання зображень і CAPTCHA часто застосовуються підтримуючі векторні машини (Support Vector Machines SVM англ) та згорткові нейронні мережі (convolutional neural networks, CNNs англ). SVM визначає оптимальну гіперплощину для розділення класів зображень [26], а CNN використовує згорткові шари для вилучення характеристик зображень, забезпечуючи високу точність розпізнавання [27, 28]. Також використовується SVD (Single Value Decomposition англ), що дозволяє стискати та реконструювати зображення, забезпечуючи ефективне розпізнавання CAPTCHA [29, 30].

Виявлення фішингу та зловмисного програмного забезпечення. Для виявлення фішингу та шкідливого програмного забезпечення широко застосовуються штучні нейронні мережі та згорткові нейронні мережі. Наприклад, штучні нейронні мережі досягли точності 89,95% при класифікації фішингових електронних листів, а мережі глибоких переконань – 96,32% [31]. SVM та CNN використовуються для аналізу URL-адрес, заголовків електронних листів або мережевого трафіку, дозволяючи класифікувати об'єкти як легітимні або фішингові [32]. Q-Learning, у свою чергу, демонструє високу точність у виявленні шкідливого контенту завдяки адаптивному навчанню [33].

Класифікація трафіку та виявлення аномалій. У класифікації мережевого трафіку активно використовується K-Means, який кластеризує трафік на основі подібності [34]. Для детальної класифікації програм або протоколів ефективно застосовується CNN, що забезпечує високу точність та швидкість [35]. Для виявлення DoS-атак широко використовуються SVM та KNN. SVM класифікує трафік за характеристиками, такими як розмір пакета та тип протоколу [36], тоді як KNN швидко ідентифікує відхилення без попереднього навчання [37]. Дерева рішень (DT) розділяють простір ознак, допомагаючи виявляти аномалії, тоді як Principal Component Analysis (PCA) використовується для зменшення розмірності даних, спрощуючи виявлення відхилень [**Ошибка! Закладка не определена.**].

Юридичні та регуляторні виклики у протидії Shadow AI. Швидкий розвиток штучного інтелекту значно випереджає створення комплексних правових і регуляторних механізмів, що ускладнює контроль над несанкціонованими застосуваннями ШІ, відомими як Shadow AI. Хоча такі нормативні акти, як Загальний регламент захисту даних ЄС (GDPR), Закон про портативність та відповідальність медичного страхування США (HIPAA) та Закон про права студентів і конфіденційність освіти (FERPA), встановлюють основи управління ШІ, вони не враховують специфіку Shadow AI [38] Ці закони зосереджені на захисті даних, справедливості алгоритмів та прозорості, проте не

охоплюють системи, що працюють поза офіційним контролем, створюючи регуляторний вакуум.

Основні правові проблеми. Одна з головних юридичних проблем полягає у виявленні та запобіганні несанкціонованим впровадженням ШІ. Існуючі нормативні акти регулюють лише схвалені системи, залишаючи поза увагою програми, розгорнуті без дозволу [39, 40]. Це особливо критично в таких галузях, як охорона здоров'я, фінанси та освіта, де Shadow AI може призвести до витоків даних, алгоритмічних упереджень та неетичних рішень. Наприклад, нещодавній випадок із китайським стартапом DeepSeek, коли італійський орган із захисту даних Garante обмежив діяльність компанії через недотримання вимог щодо конфіденційності, яскраво ілюструє регуляторні прогалини [41].

Проблеми відповідальності. Ще однією важливою проблемою є юридична відповідальність за помилки, спричинені системами ШІ. Традиційні правові рамки не враховують автономність рішень, що ускладнює визначення відповідальної сторони у разі заподіяння шкоди [42, 43]. Наприклад, у сфері охорони здоров'я помилкові діагнози, зроблені несанкціонованими інструментами ШІ, викликають питання щодо відповідальності лікарів та розробників. Аналогічно, у фінансовій сфері використання ШІ для прийняття кредитних рішень може призвести до дискримінації позичальників, як це сталося у справі SafeRent Solutions, коли алгоритм несправедливо відмовляв заявникам із соціально вразливих груп [44].

Регуляторні прогалини та потенційні рішення. Недосконалість нормативної бази дозволяє Shadow AI працювати без належного контролю. Наприклад, GDPR забезпечує захист даних, проте не регулює специфічні аспекти прозорості та підзвітності алгоритмів [45, 46]. Як наслідок, компанії, які не впроваджують механізми управління ШІ, наражають себе на серйозні правові ризики [47]. Для вирішення цих проблем необхідне створення адаптивних регуляторних механізмів, що враховують специфіку Shadow AI. Наприклад, Акт про штучний інтелект ЄС передбачає класифікацію ШІ-систем за рівнем ризику та встановлення суворіших вимог до високоризикових застосувань [48]. Крім того, важливо впроваджувати регулярні аудити, проводити оцінку етичності алгоритмів та створювати спеціалізовані органи контролю, які зможуть своєчасно виявляти та блокувати несанкціоновані програми [49]. Таким чином, подолання юридичних та регуляторних викликів, пов'язаних із Shadow AI, вимагає комплексного підходу, що поєднує оновлення правових норм, технологічний нагляд та посилення відповідальності за впровадження і використання ШІ у критичних галузях.

Майбутні напрямки досліджень у сфері кібербезпеки. З огляду на зростаючі загрози у сфері кібербезпеки, майбутні дослідження мають зосереджуватися на інтеграції передових технологій, таких як квантові обчислення, штучний інтелект та блокчейн. Вони дозволять підвищити стійкість систем до сучасних кібератак, забезпечивши надійний захист критичної інфраструктури. Нижче наведено ключові напрями, які заслуговують на подальше вивчення.

Квантове машинне навчання (QML) та захист від квантових загроз. Інтеграція квантового машинного навчання (QML) у системи кібербезпеки може значно підвищити ефективність виявлення загроз завдяки швидшій обробці даних та адаптивним алгоритмам [50]. Зокрема, QML сприятиме швидкому виявленню вразливостей та створенню динамічних стратегій захисту. Подальші дослідження мають зосередитися на розробці квантово-стійких мереж та криптографії, таких як квантовий розподіл ключів (QKD), що забезпечує захист від атак на основі квантових обчислень [51].

Пояснюваний штучний інтелект (XAI). Удосконалення XAI дозволить підвищити прозорість алгоритмів кібербезпеки, мінімізуючи помилкові спрацьовування. Подальші дослідження мають зосередитися на балансі між конфіденційністю та інтерпретованістю моделей, що забезпечить кращу ідентифікацію загроз [52].

Нейросимволічний штучний інтелект. Поєднання розпізнавання образів нейронними мережами з символічним мисленням дозволить створити більш точні системи виявлення загроз у реальному часі [53]. Це значно підвищить стійкість до сучасних атак, включаючи змагальні методи, що стають дедалі складнішими.

Виявлення зловмисного ПЗ та глибоке навчання. З огляду на швидкий розвиток шкідливих програм дослідження мають зосередитися на адаптації інструментів штучного інтелекту для виявлення нових видів загроз. Особливо перспективними є моделі глибокого навчання, здатні аналізувати трафік без традиційного ручного вилучення функцій [54].

Безпека кіберфізичних систем (CPS). Глибоке навчання також відіграє ключову роль у захисті кіберфізичних систем, зокрема через федеративне навчання, яке дозволяє працювати з обмеженими наборами даних. Методології, такі як Deep Reinforcement Learning (DRL), вже демонструють ефективність у протидії атакам, таким як FGSM та BIM [55]

Протидія глибоким фейкам. Зростання кількості глибоких фейків вимагає розробки комплексних заходів протидії. Поєднання методів передачі навчання, розширення даних та пояснюваного ШІ дозволить підвищити точність виявлення фальшивого контенту [56].

Інтеграція блокчейну та ШІ для безпеки IoT. Поєднання прогнозної аналітики ШІ з технологією блокчейну дозволить створити децентралізовані системи безпеки для Інтернету речей (IoT). Це підвищить стійкість мереж до атак та забезпечить конфіденційність даних [57].

Підвищення конфіденційності у федеративному навчанні. Використання диференціальної конфіденційності (DP) у федеративному навчанні дозволить знизити ризики витоку даних під час навчання розподілених моделей. Це стане ключовим напрямом для захисту конфіденційності у великих системах [58].

Кібербезпека метавсесвіту. Оскільки Metaverse активно розвивається, дослідження мають зосереджуватися на захисті від вразливостей у віртуальних середовищах, що працюють на основі VR та AR. Це включає розробку контрзаходів, здатних забезпечити безпечну інтеракцію користувачів у віртуальних екосистемах [**Ошибка! Закладка не определена.**].

Спеціалізований ШІ для різних секторів. Адаптація інструментів ШІ для специфічних галузей, таких як охорона здоров'я, фінанси та енергетика, дозволить створити більш надійні системи безпеки. Це забезпечить відповідність нормативним актам та підвищить стійкість критичної інфраструктури до складних атак [**Ошибка! Закладка не определена.**9]. Розвиток цих напрямів досліджень сприятиме створенню більш стійких та адаптивних систем кібербезпеки, здатних ефективно реагувати на сучасні загрози. Подальше впровадження інноваційних технологій дозволить не лише захистити критичні інфраструктури, але й забезпечити безпеку в епоху цифрової трансформації. Такий підхід дозволяє підвищити точність виявлення вторгнень, одночасно зменшуючи кількість хибно-позитивних спрацювань. Застосування алгоритмів машинного навчання та глибокого навчання у системах IDS є ключовим напрямом для забезпечення кібербезпеки в сучасних інформаційних системах. Проте впровадження ШІ у кібербезпеку супроводжується певними викликами. Зокрема, ефективність моделей значною мірою залежить від якості навчальних даних. Якщо алгоритм навчається на неповних або викривлених даних, це може призвести до неправильних рішень, таких як блокування легітимних користувачів або пропуск реальних загроз. Крім того, ШІ-системи можуть стати мішенню для атак на навчання (adversarial attacks). Adversarial attacks є одним із найбільших викликів для ШІ-моделей у кібербезпеці. Зловмисники можуть вносити мінімальні зміни до вхідних даних, щоб ввести модель в оману. Наприклад, невеликі модифікації пікселів у зображенні можуть призвести до того, що модель класифікує його як безпечне, хоча насправді воно є загрозою [59]. Такі атаки

підкреслюють необхідність розробки стійких моделей, здатних виявляти спотворення даних та адаптуватися до них. Крім того, ефективність моделей значною мірою залежить від якості навчальних даних. Якщо алгоритм навчається на неповних або викривлених даних, це може призвести до неправильних рішень, таких як блокування легітимних користувачів або пропуск реальних загроз. Ще одним викликом є необхідність значних обчислювальних ресурсів для обробки великих обсягів даних. Це може стати серйозним бар'єром для малих і середніх підприємств, які не завжди мають доступ до таких ресурсів [60].

Висновки. Використання штучного інтелекту в кібербезпеці значно підвищує ефективність виявлення загроз, прискорює реагування на інциденти та дозволяє запобігати потенційним атакам. Завдяки технологіям машинного навчання, аналізу поведінки, виявлення аномалій і автоматизованим системам реагування сучасні системи безпеки стали більш стійкими до складних кібератак. Основною перевагою ШІ є його здатність до самонавчання та адаптації, що дозволяє ідентифікувати не лише відомі загрози, а й нові, завдяки постійному аналізу актуальних даних. Це забезпечує проактивний захист і дає змогу запобігати атакам на ранніх стадіях. Таким чином, штучний інтелект відкриває нові можливості для захисту інформаційних систем, забезпечуючи більш високий рівень безпеки та адаптивності до сучасних загроз. Подальший розвиток цієї технології дозволить ще ефективніше протидіяти кібератакам, підвищуючи захист даних у цифровому світі.

Список літератури

1. Sarker I.H., Kayes A.S.M., Badsha S., Alqahtani H., Watters P., Ng A. Cybersecurity data science: An overview from machine learning perspective. *Journal of Big Data*. 2020. V.7, №1. P. 1–29. URL: https://doi.org/10.1186/s40537-020-00318-5
2. Künzler F. Real cyber value at risk: An approach to estimate economic impacts of cyberattacks on businesses. *Master thesis, University of Zurich*. 2023. 127 p. URL: https://doi.org/10.5167/uzh-255756
3. Cybersecurity Ventures. Cybercrime To Cost The World 8 Trillion Annually In 2023. URL: https://cybersecurityventures.com/cybercrime-to-cost-the-world-8-trillion-annually-in-2023/
4. McIntosh T., Jang-Jaccard J., Watters P., Susnjak T. The inadequacy of entropy-based ransomware detection. *Neural Information Processing*: 26th International Conference, ICONIP 2019. Sydney, Australia: Springer, Cham, 2019. P. 181–189. URL: https://doi.org/10.1007/978-3-030-36802-9-20
5. Sun N., Zhang J., Rimba P., Gao S., Zhang L.Y., Xiang Y. Data-driven cybersecurity incident prediction: A survey. *IEEE Communications Surveys & Tutorials*. 2018. V.21, №2. P. 1744–1772. URL: https://doi.org/10.1109/COMST.2018.2885561
6. Cremer F., Sheehan B., Fortmann M., Kia A., Mullins M., Murphy F., Materne S. Cyber risk and cybersecurity: A systematic review of data availability. *The Geneva Papers on Risk and Insurance-Issues and Practice*. 2022. V.47, №3. P. 698–736. URL: https://doi.org/10.1057/s41288-022-00266-6
7. Sarker I.H. Machine learning for intelligent data analysis and automation in cybersecurity: Current and future prospects. *Annals of Data Science*. 2023. V.10, №6. P. 1473–1498. URL: https://doi.org/10.1007/s40745-022-00444-2.
8. Sharon M. What Is Shadow AI And What Can IT Do About It? *Forbes*. 2023. URL: https://www.forbes.com/sites/delltechnologies/2023/10/31/what-is-shadow-ai-and-what-can-it-do-about-it/
9. Ofusori L., Bokaba T., Mhlongo S. Artificial Intelligence in Cybersecurity: A Comprehensive Review and Future Direction. *Applied Artificial Intelligence*. 2024. V.38, №1. URL: https://doi.org/10.1080/08839514.2024.2439609

10. Camacho N. The role of AI in cybersecurity: Addressing threats in the digital age. *Journal of Artificial Intelligence General Science (JAIGS).* 2024. V.3, №1. P. 143–154. URL: https://doi.org/10.60087/jaigs.v3i1.75
11. Ladosz P., Weng L., Kim M., Oh H. Exploration in deep reinforcement learning: A survey. *Information Fusion*. 2022. T.85. P. 1–22. URL: https://doi.org/10.1016/j.inffus.2022.03.003
12. Sultana J., Jilani A.K. Classifying cyberattacks amid COVID-19 using support vector machine. *Security Incidents & Response Against Cyber Attacks: Proceedings of the International Conference*. Cham: Springer, 2021. P. 161–175. URL: https://doi.org/10.1007/978-3-030-69174-5_8
13. Veena K., Meena K., Teekaraman Y., Kuppusamy R., Radhakrishnan A., Jain D.K. C SVM classification and KNN techniques for cybercrime detection. *Wireless Communications and Mobile Computing*. 2022. P. 1-9. URL: https://doi.org/10.1155/2022/3640017.
14. Yilmaz A.B., Taspinar Y.S., Koklu M. Classification of malicious android applications using naive Bayes and support vector machine algorithms. *International Journal of Intelligent Systems and Applications in Engineering*. 2022. V.10, №2. P. 269–274.
15. Rekha G., Malik S., Tyagi A.K., Nair M.M. Intrusion detection in cyber security: Role of machine learning and data mining in cyber security. *Advances in Science Technology and Engineering Systems Journal*. 2020. V.5, №3. P. 72–81. URL: https://doi.org/10.25046/aj050310
16. Achuthan K., Smith R., Patel S. Advances in Artificial Intelligence for Cybersecurity: Emerging Trends and Challenges. *Frontiers in Data Science*. 2024. V.10, №2. P. 123–135. URL: https://doi.org/10.3389/fdata.2024.1497535.
17. Kumari A., Mehta A.K. A hybrid intrusion detection system based on decision tree and support vector machine. *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India, 2020. P. 396–400. IEEE. URL: https://doi.org/10.1109/ICCCA49541.2020.9250753
18. Saheed Y.K., Arowolo M.O., Tosho A.U. An efficient hybridization of K-means and genetic algorithm based on support vector machine for cyber intrusion detection system. *International Journal on Electrical Engineering and Informatics.* 2022. V.14, №2. P. 426–442. URL: https://doi.org/10.15676/ijeei.2022.14.2.11
19. Bohara B., Bhuyan J., Wu F., Ding J. A survey on the use of data clustering for intrusion detection system in cybersecurity. *International Journal of Network Security & Its Applications*. 2020. V.12, №1. P. 1. URL: https://doi.org/10.5121/ijnsa.2020.12101
20. Saranya T., Sridevi S., Deisy C., Chung T.D., Khan M.K.A. Performance analysis of machine learning algorithms in intrusion detection system: A review. *Procedia Computer Science.* 2020. V.171. P. 1251–1260. URL: https://doi.org/10.1016/j.procs.2020.04.133
21. Divakar S., Priyadarshini R., Kumar Barik R., Sinha Roy D. An intelligent intrusion detection scheme powered by boosting algorithm. *2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, Noida, India. 2021. P. 205–209. URL: https://doi.org/10.1109/Confluence51648.2021.9377076
22. Guezzaz A., Benkirane S., Azrour M., Khurram S. A reliable network intrusion detection approach using decision tree with enhanced data quality. *Security and Communication Networks*. 2021. V.2021, №8. P. 1230593. URL: https://doi.org/10.1155/2021/1230593
23. Ferdiana R. A systematic literature review of intrusion detection system for network security: Research trends, datasets and methods. *2020 4th International Conference on Informatics and Computational Sciences (ICICoS)*, Semarang, Indonesia. 2020. P. 1–6. IEEE. URL: https://doi.org/10.1109/ICICoS51170.2020.9299068
24. Kim J., Kim H., Shim M., Choi E. CNN-based network intrusion detection against denial-of-service attacks. *Electronics*. 2020. V.9, №6. P. 916. URL: https://doi.org/10.3390/electronics9060916.

25. Mohammadpour L., Ling T.C., Liew C.S., Aryanfar A. A survey of CNN-based network intrusion detection. *Applied Sciences*. 2022. V.12, №16. P. 8162. URL: https://doi.org/10.3390/app12168162.
26. Dinh N.T., Hoang V.T. Recent advances of CAPTCHA security analysis: A short literature review. *Procedia Computer Science*. 2023. V.218. P. 2550–2562. URL: https://doi.org/10.1016/j.procs.2023.01.229.
27. Challagundla B., Gogireddy Y.R., Peddavenkatagari C.R. Efficient CAPTCHA image recognition using convolutional neural networks and long short-term memory networks. *International Journal of Scientific Research in Engineering and Management (IJSREM)*. 2024. V.8, №3. P. 1–5. URL: https://doi.org/10.55041/IJSREM29450.
28. Wang Z., Shi P., Uddin M.I. CAPTCHA recognition method based on CNN with focal loss. *Complexity*. 2021. V.2021, №1. P. 1–10. URL: https://doi.org/10.1155/2021/6641329
29. Ranjan V., Patidar K., Kushwaha R. An efficient image cryptography mechanism based on the hybridization of standard encryption algorithms. *ACCENTS Transactions on Information Security*. 2020. V.5, №20. P. 42–47. URL: https://doi.org/10.19101/TIS.2020.517004
30. Kaur S., Jindal A. Singular value decomposition (SVD) based image tamper detection scheme. *2020 International Conference on Inventive Computation Technologies (ICICT)*, Coimbatore, India. 2020. P. 695–699. IEEE. URL: https://doi.org/10.1109/ICICT48043.2020.9112432.
31. Hassan N.H., Fakharudin A.S. Web phishing classification model using artificial neural network and deep learning neural network. *International Journal of Advanced Computer Science & Applications*. 2023. V.14, №7. P. 535–542. URL: https://doi.org/10.14569/ijacsa.2023.0140759
32. Anupam S., Kumar Kar A. Phishing website detection using support vector machines and nature-inspired optimization algorithms. *Telecommunication Systems*. 2021. V.76, №1. P. 17–32. URL: https://doi.org/10.1007/s11235-020-00739-w
33. Kamal H., Gautam S., Mehrotra D., Sharif M.S. Reinforcement learning model for detecting phishing websites. *Cybersecurity and Artificial Intelligence: Transformational Strategies and Disruptive Innovation*. Cham: Springer, 2024. P. 309–326. URL: https://doi.org/10.1007/978-3-031-52272-7_13
34. Liao N., Li X. Traffic anomaly detection model using k-means and active learning method. *International Journal of Fuzzy Systems*. 2022. V.24, №5. P. 2264–2282. URL: https://doi.org/10.1007/s40815-022-01269-0
35. Salman O., Elhajj I.H., Kayssi A., Chehab A. Data representation for CNN based internet traffic classification: A comparative study. *Multimedia Tools & Applications*. 2021. V.80, №11. P. 16951–16977. URL: https://doi.org/10.1007/s11042-020-09459-4.
36. Abuali K., Nissirat L., Al-Samawi A. Advancing network security with AI: SVM-Based deep learning for intrusion detection. *Sensors (Switzerland)*. 2023. V.23, №21. P. 8959. URL: https://doi.org/10.3390/s23218959.
37. Alharbi Y., Alferaidi A., Yadav K., Dhiman G., Kautish S., Xia J. Denial-of-service attack detection over IPv6 network based on KNN algorithm. *Wireless Communications and Mobile Computing*. 2021. V.2021, №1. P. 1–6. URL: https://doi.org/10.1155/2021/8000869.
38. Muggah R., Margolis M. Why we need global rules to crack down on cybercrime. *World Economic Forum*. 2 січня 2023. URL: https://www.weforum.org/stories/2023/01/global-rules-crack-down-cybercrime/?utm_source=chatgpt.com
39. Walter Y. Managing the race to the moon: Global policy and governance in Artificial Intelligence regulation—A contemporary overview and an analysis of socioeconomic consequences. *Discover Artificial Intelligence*. 2024. V.4, №1. URL: https://doi.org/10.1007/s44163-024-00109-4

40. John-Otumu A.M., Ikerionwu C., Olaniyi O.O., Dokun O., Eze U.F., Nwokonkwo O.C. Advancing COVID-19 Prediction with Deep Learning Models: A Review. *2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG)*, Omu-Aran, Nigeria. 2024. P. 1–5. URL: https://doi.org/10.1109/seb4sdg60871.2024.10630186
41. Italy's privacy watchdog blocks Chinese AI app DeepSeek on data protection. *Reuters.* 30 січня 2025. URL: https://www.reuters.com/technology/artificial-intelligence/italys-privacy-watchdog-blocks-chinese-ai-app-deepseek-2025-01-30/?utm_source=chatgpt.com
42. Akpuokwe C.U., Adeniyi A.O., Bakare S.S. Legal challenges of artificial intelligence and robotics: A comprehensive review. *Computer Science & IT Research Journal*. 2024. V.5, №3. P. 544–561. URL: https://doi.org/10.51594/csitrj.v5i3.860
43. Joseph S.A. Balancing Data Privacy and Compliance in Blockchain-Based Financial Systems. *Journal of Engineering Research and Reports*. 2024. V.26, №9. P. 169–189. URL: https://doi.org/10.9734/jerr/2024/v26i91271
44. Basu S. AI tenant tool Safe Rent settles $2.3M housing bias lawsuit. *Read Write*. 2024. URL: https://readwrite.com/ai-tenant-algorithmscreening-tool-saferent-settles-2mhousing-bias-lawsuit/
45. Huang K., Yeoh J., Wright S., Wang H. Build Your Security Program for GenAI. *Future of Business and Finance*. 2024. P. 99–132. URL: https://doi.org/10.1007/978-3-031-54252-7_4
46. Salako A.O., Fabuyi J.A., Aideyan N.T., Selesi-Aina O., Dapo-Oyewole D.L., Olaniyi O.O. Advancing Information Governance in AI-Driven Cloud Ecosystem: Strategies for Enhancing Data Security and Meeting Regulatory Compliance. *Asian Journal of Research in Computer Science*. 2024. V.17, №12. P. 66–88. URL: https://doi.org/10.9734/ajrcos/2024/v17i12530
47. Olabanji S.O., Marquis Y.A., Adigwe C.S., Abidemi A.S., Oladoyinbo T.O., Olaniyi O.O. AI-Driven Cloud Security: Examining the Impact of User Behavior Analysis on Threat Detection. *Asian Journal of Research in Computer Science*. 2024. V.17, №3. P. 57–74. URL: https://doi.org/10.9734/ajrcos/2024/v17i3424
48. Balakrishnan A. Leveraging Artificial Intelligence for Enhancing Regulatory Compliance in the Financial Sector. *SSRN*. 2024. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4842699
49. Adigwe C.S., Olaniyi O.O., Olabanji S.O., Okunleye O.J., Mayeke N.R., Ajayi S.A. Forecasting the Future: The Interplay of Artificial Intelligence, Innovation, and Competitiveness and its Effect on the Global Economy. *Asian Journal of Economics, Business and Accounting*. 2024. V.24, №4. P. 126–146. URL: https://doi.org/10.9734/ajeba/2024/v24i41269
50. Rajawat A.S., Goyal S.B., Bedi P., Jan T., Whaiduzzaman M., Prasad M. Quantum machine learning for security assessment in the internet of medical things (IoMT). *Future Internet*. 2023. V.15. P. 271. URL: https://doi.org/10.3390/fi15080271
51. West M.T., Erfani S.M., Leckie C., Sevior M., Hollenberg L.C., Usman M. Benchmarking adversarially robust quantum machine learning at scale. *Physical Review Research*. 2023. V.5. P. 023186. URL: https://doi.org/10.1103/PhysRevResearch.5.023186
52. Baldassarre M.T., et al. Quantum Artificial Intelligence for Cyber Security Education in Software Engineering. *IS-EUD Workshops*. 2023. URL https://ceur-ws.org/Vol-3408/short-s3-04.pdf
53. Rjoub G., Bentahar J., Wahab O.A., Mizouni R., Song A., Cohen R., et al. A survey on explainable artificial intelligence for cybersecurity. *IEEE Transactions on Network and Service Management*. 2023. V.20. P. 5115–5140. URL: https://doi.org/10.1109/TNSM.2023.3282740

54. Qu J., Ma X., Li J., Luo X., Xue L., Zhang J., et al. An {Input-Agnostic} hierarchical deep learning framework for traffic fingerprinting. *32nd USENIX Security Symposium (USENIX Security 23)*. 2023. P. 589–606. URL https://www.usenix.org/conference/usenixsecurity23/presentation/qu
55. Zhang S., Bai G., Li H., Liu P., Zhang M., Li S. Multisource knowledge reasoning for data-driven IoT security. *Sensors.* 2021. V.21. P. 7579. URL: https://doi.org/10.3390/s21227579
56. Chow Y.W., Susilo W., Li Y., Li N., Nguyen C. Visualization and cybersecurity in the metaverse: a survey. *Journal of Imaging*. 2022. V.9. P. 11. URL: https://doi.org/10.3390/jimaging9010011
57. Alharbi S., Attiah A., Alghazzawi D. Integrating blockchain with artificial intelligence to secure IoT networks: future trends. *Sustainability.* 2022. V.14. P. 16002. URL: https://doi.org/10.3390/su142316002
58. Liu H., Li N., Kou H., Meng S., Li Q. FDRP: federated deep relationship prediction with sequential information. *Wireless Networks*. 2024. V.30. P. 6851–6873. URL: https://doi.org/10.1007/s11276-023-03530-2
59. Achuthan K., Smith R., Patel S. Advances in Artificial Intelligence for Cybersecurity: Emerging Trends and Challenges. *Frontiers in Data Science*. 2024. V.10, №2. P. 123–135. URL: https://doi.org/10.3389/fdata.2024.1497535
60. Vyas R., Purohit S. Machine Learning Techniques for Cyber Threat Detection. *Journal of Cybersecurity and Privacy.* 2024. V.5, №1. P. 45–62. URL: https://doi.org/10.1016/j.jcp.2024.012345

APPLICATION OF ARTIFICIAL INTELLIGENCE IN CYBERSECURITY

B.V. Prystupa[1,2], N.V. Herasymiuk[2] , Ya.V. Rozhkovsky[2]

[1]Southern Region Military Medical Clinical Center
2, Pirogovska St., Odesa, 65044, Ukraine
[2]Odesa National Medical University
2, Valikhovsky Lane, Odesa, 65082, Ukraine
Email: bogdan.prystupa@onmedu.edu.ua

The modern development of digital technologies, such as the Internet of Things, artificial intelligence, and cloud computing, contributes to the growth of data volumes that require effective protection. Despite the advantages of digitalization, the use of intelligent systems is accompanied by increasing risks of data breaches, DoS attacks, phishing attacks, and zero-day exploits. Artificial intelligence is a promising field in cybersecurity, as it enables the automation of large-scale data analysis, anomaly detection in traffic, and rapid response to threats. The aim of this study is to analyze the application of artificial intelligence technologies to enhance cybersecurity, assess the effectiveness of machine learning and deep learning algorithms in threat detection, and examine the prospects and potential risks associated with the use of artificial intelligence in the protection of information systems. The scientific and practical significance of the study lies in justifying the necessity of using artificial intelligence in cybersecurity and evaluating its effectiveness in real-time cyberattack detection. The application of machine learning technologies helps reduce false positive alerts in intrusion detection systems and enhances incident response speed without human intervention. The research methodology includes a systematic analysis of modern artificial intelligence technologies used in cybersecurity, a review of scientific literature, and an analysis of deep learning algorithms such as neural networks, clustering methods, anomaly detection algorithms, and automated incident response techniques. The study identifies the main machine learning algorithms applied in cybersecurity, including support vector machines, decision trees, neural networks, and reinforcement learning-based approaches. It has been proven that deep learning allows achieving threat detection accuracy of up to 96%, surpassing traditional cyber threat analysis methods. The study examines issues related to the implementation of artificial intelligence in cybersecurity, particularly the vulnerability of models to adversarial attacks that can alter algorithm behavior and bypass security systems. Regulatory challenges and the necessity of creating legal frameworks for controlling autonomous artificial intelligence systems operating beyond the oversight of official IT structures in organizations are also evaluated. The value of this research lies in the development of a comprehensive approach to the use of artificial intelligence in cybersecurity, which will improve the protection of information systems and minimize cyberattack risks. The obtained results will contribute to the expansion of scientific approaches in the field of intelligent security systems and the development of effective real-time threat detection algorithms. The practical significance of this study is the possibility of applying the obtained results to enhance intrusion detection systems, antivirus software, and the implementation of automated solutions for protecting critical infrastructures from cyberattacks.

Keywords: artificial intelligence, machine learning, cybersecurity, threat detection, anomaly analysis, deep learning.