НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ОДЕСЬКА ЮРИДИЧНА АКАДЕМІЯ»
ФАКУЛЬТЕТ КІБЕРБЕЗПЕКИ ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Кафедра кібербезпеки

КВАЛІФІКАЦІЙНА РОБОТА

на тему:

«ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ У СФЕРІ КІБЕРБЕЗПЕКИ»

Виконала: здобувачка 4 курсу
Рівень бакалавр
Галузь знань 12 «Інформаційні технології»,
спеціальність 125 «Кібербезпека»
Савескул Анастасія Дмитрівна

Керівник: к.т.н., доцент
Кухаренко Сергій Вікторович

Рецензент: д.т.н., професор
Соколов Артем Вікторович

Одеса – 2025

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ОДЕСЬКА ЮРИДИЧНА АКАДЕМІЯ»
Факультет **кібербезпеки та інформаційних технологій**
Кафедра **кібербезпеки**
Галузь знань: **12 Інформаційні технології**
Спеціальність: **125 Кібербезпека**
Назва освітньої програми: **Кібербезпека**

ЗАТВЕРДЖУЮ
Завідувач кафедри кібербезпеки
професор С.А. Горбаченко
_____ «____ »_____ 20__ року

З А В Д А Н Н Я
на кваліфікаційну (бакалаврську) роботу
здобувачці Савескул Анастасії Дмитрівні

1. Тема роботи: «Використання штучного інтелекту у сфері кібербезпеки»
Керівник роботи: к.т.н, доцент Кухаренко Сергій Вікторович
затверджені наказом НУ ОЮА від 18 березня 2025 року № 548-6
2. Строк подання здобувачем роботи 19 травня 2025 року
3. Дата видачі завдання 16 вересня 2024 року

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів бакалаврської роботи	Строк виконання етапів роботи	Примітка
1	Аналіз предметної області та пошук науково-технічної інформації за темою роботи	Вересень – жовтень 2024	
2	Узгодження теми з науковим керівником, формулювання мети завдань дослідження	Листопад 2024	
3	Робота над першим розділом	Листопад – грудень 2024	
4	Робота над другим розділом	Січень – лютий 2025	
5	Робота над третім розділом	Лютий – березень 2025	
6	Уточнення подальшого обсягу роботи та його коригування	Березень – травень 2025	
7	Оформлення звітної частини	Травень 2025	
8	Захист роботи	10.06.2025	

Здобувач _____
(підпис) (прізвище та ініціали)
Керівник роботи _____
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Кваліфікаційна робота на тему «Використання штучного інтелекту у сфері кібербезпеки» на здобуття першого (бакалаврського) рівня вищої освіти за спеціальністю 125 Кібербезпека, освітньою програмою: Кібербезпека, містить 4 рисунки, 17 літературних джерел за переліком посилань. Робота виконана на 41 сторінці загального тексту та 34 сторінках основного тексту.

Метою роботи є аналіз та дослідження застосування методів штучного інтелекту у сфері кібербезпеки, включаючи машинне навчання для виявлення аномалій, нейронні мережі для аналізу загроз, автоматизацію процесів кіберзахисту та протидію атакам за допомогою штучного інтелекту.

У ході дослідження були розглянуті методи машинного навчання для виявлення шкідливих програм, використання нейронних мереж для виявлення аномальних дій у мережевому трафіку та алгоритми прогнозування кіберзагроз. Також було проведено аналіз ефективності різних підходів та запропоновані рекомендації щодо підвищення рівня безпеки за допомогою технологій штучного інтелекту.

Можливі напрями подальших досліджень включають розробку нових методів глибокого навчання для кіберзахисту, вдосконалення алгоритмів адаптивного виявлення загроз та інтеграцію штучного інтелекту у системи управління інформаційною безпекою.

КІБЕРБЕЗПЕКА, ШТУЧНИЙ ІНТЕЛЕКТ, МАШИННЕ НАВЧАННЯ, АНАЛІЗ ЗАГРОЗ, АВТОМАТИЗАЦІЯ ЗАХИСТУ, ВИЯВЛЕННЯ АНОМАЛІЙ.

ABSTRACT

The qualification work on the topic "Application of Artificial Intelligence in Cybersecurity" for obtaining the first (bachelor's) level of higher education in specialty 125 Cybersecurity**,** educational program: Cybersecurity, contains 4 figures, and 17 literary sources in the reference list. The work consists of 41 pages of total text and 34 pages of main text.

The purpose of the work is to analyze and study the application of artificial intelligence methods in the field of cybersecurity, including machine learning for anomaly detection, neural networks for threat analysis, automation of cybersecurity processes, and countering attacks using artificial intelligence.

During the research, machine learning methods for detecting malware, the use of neural networks to identify anomalous activities in network traffic, and algorithms for predicting cyber threats were examined. An analysis of the effectiveness of various approaches was conducted, and recommendations were proposed to enhance security levels using artificial intelligence technologies.

Possible directions for further research include the development of new deep learning methods for cybersecurity, improvement of adaptive threat detection algorithms, and the integration of artificial intelligence into information security management systems.

CYBERSECURITY, ARTIFICIAL INTELLIGENCE, MACHINE LEARNING, THREAT ANALYSIS, SECURITY AUTOMATION, ANOMALY DETECTION

ЗМІСТ

ВСТУП

Сучасний світ стрімко розвивається завдяки цифровим технологіям, що стали невід'ємною частиною повсякденного життя. Інформаційні системи, великі обсяги даних та глобальна мережа Інтернет створюють нові можливості для комунікації, бізнесу та науки. Водночас із цими можливостями зростає і кількість загроз, пов'язаних із кібербезпекою. Зловмисники використовують новітні технології для проведення атак, спрямованих на викрадення даних, компрометацію інформаційних систем і навіть масштабні кібертерористичні акти. У зв'язку з цим розробка нових методів захисту інформації стає однією з ключових проблем сучасного суспільства.

Одним із перспективних напрямів вирішення цієї проблеми є використання штучного інтелекту (ШІ) у сфері кібербезпеки. ШІ дозволяє аналізувати величезні обсяги даних, автоматизувати процеси виявлення загроз та швидко реагувати на нові атаки [1]. Технології машинного навчання, нейронні мережі та алгоритми глибокого навчання значно підвищують ефективність систем кіберзахисту, роблячи їх більш гнучкими та адаптивними до сучасних викликів.

Застосування штучного інтелекту у сфері кібербезпеки включає широкий спектр методів і технологій. Це, зокрема, аналіз поведінкових аномалій у мережевому трафіку, виявлення шкідливого програмного забезпечення, запобігання фішинговим атакам та боротьба з соціальною інженерією. Використовуючи передові алгоритми, ШІ може прогнозувати потенційні атаки, зменшуючи ризики та підвищуючи рівень безпеки інформаційних систем.

Проте впровадження штучного інтелекту в кібербезпеку має і свої виклики. Це, зокрема, проблема пояснюваності рішень, зроблених алгоритмами, ризики помилкових спрацьовувань, а також можливість використання ШІ для створення нових, ще складніших видів атак. Тому дослідження даної теми є надзвичайно актуальним та має велике значення для подальшого розвитку методів кіберзахисту.

Об'єктом дослідження є процеси забезпечення кібербезпеки в умовах стрімкого розвитку цифрових технологій.

Предметом дослідження є методи та алгоритми застосування штучного інтелекту для виявлення, запобігання та реагування на кіберзагрози. Тому дослідження даної теми є надзвичайно актуальним та має велике значення для подальшого розвитку методів кіберзахисту.

Метою даної роботи є дослідження можливостей застосування штучного інтелекту у сфері кібербезпеки, аналіз сучасних методів та алгоритмів, що використовуються для виявлення та запобігання кіберзагрозам, а також оцінка їх ефективності.

У ході дослідження будуть розглянуті практичні аспекти застосування машинного навчання, нейронних мереж та автоматизованих систем кіберзахисту, що дозволить зробити висновки про перспективи розвитку даного напряму. Таким чином, дана робота спрямована на виявлення сильних та слабких сторін існуючих методів, розробку рекомендацій щодо покращення кібербезпеки за допомогою ШІ та визначення перспектив подальших досліджень у цій галузі.

1 АНАЛІЗ СУЧАСНИХ ЗАГРОЗ ТА МЕТОДІВ КІБЕРЗАХИСТУ

1.1 Основні загрози кібербезпеці у сучасному цифровому середовищі

У сучасному цифровому середовищі питання кібербезпеки набуває особливої актуальності через стрімке зростання використання інформаційних технологій у всіх сферах життєдіяльності людини та суспільства. Від приватних користувачів до міжнародних корпорацій – усі стикаються з ризиками, що виникають унаслідок кіберзагроз. Ці загрози набули не лише масштабного характеру, а й вражають своєю різноманітністю та складністю [1].

Класифікацію основних кіберзагроз наведено у рисунку 1.1.

Рисунок 1.1 – Класифікація основних кіберзагроз

Шкідливе програмне забезпечення (Malware) є одним із найпоширеніших інструментів кіберзлочинців. До malware відносяться віруси, трояни, хробаки, програми-вимагачі (ransomware) та шпигунські програми. Зловмисники використовують шкідливе ПЗ для викрадення даних, блокування доступу до систем або шифрування файлів з подальшою вимогою викупу за їх відновлення. За даними

звітів компанії Statista, у 2023 році кількість атак програмами-вимагачами зросла на 80% порівняно з попереднім роком.

Фішинг та соціальна інженерія базуються на обмані користувачів для отримання їхніх особистих або фінансових даних. Зловмисники часто надсилають фальшиві листи нібито від банків чи авторитетних організацій, змушуючи жертву перейти за шкідливим посиланням або надати свої облікові дані. У 2022 році згідно з даними Verizon Data Breach Investigations Report, 36% усіх кіберінцидентів були спричинені саме фішингом.

Розподілені атаки на відмову в обслуговуванні (DDoS-атаки) призводять до недоступності вебсайтів або сервісів шляхом масового перевантаження їх ресурсів. DDoS-атаки завдають значних фінансових збитків бізнесу: за підрахунками IBM, середня вартість простою великої компанії внаслідок DDoS-атаки сягає понад 200 000 доларів за годину.

Злом облікових записів та витік даних стали серйозною проблемою для багатьох організацій. Компрометація паролів через вразливості систем або фішингові атаки призводить до крадіжки особистих даних користувачів, банківської інформації або конфіденційної інформації компаній. Прикладом є атака на сервіс Dropbox у 2022 році, унаслідок якої було скомпрометовано електронні адреси та паролі мільйонів користувачів.

Вразливості програмного забезпечення та апаратних систем залишаються постійною ціллю для кіберзлочинців. Помилки у коді або недоліки в архітектурі пристроїв (наприклад, вразливість Log4Shell у 2021 році) відкривають шлях до несанкціонованого доступу або повного контролю над системами.

Кібертероризм та державні атаки є особливо небезпечним явищем. Використання кіберзасобів для паралізації критичної інфраструктури, маніпуляції громадською думкою або втручання у виборчі процеси становить реальну загрозу для національної безпеки.

Атаки на Інтернет речей (IoT) стали масовим явищем через недостатню захищеність підключених пристроїв. Камери спостереження, розумні телевізори,

побутова техніка часто стають точкою входу для атак на корпоративні мережі або приватні дані.

Кожен із вказаних типів атак демонструє, що сучасне цифрове середовище є вкрай вразливим перед обличчям швидко еволюціонуючих загроз. Відповідно, потреба у вдосконаленні засобів захисту, зокрема із застосуванням штучного інтелекту, є критично важливою для ефективної протидії кібератакам.

1.2 Традиційні методи захисту інформаційних систем

У відповідь на зростання кіберзагроз було розроблено широкий спектр традиційних методів захисту інформаційних систем. Ці заходи охоплюють як технічні рішення, так і організаційні політики, спрямовані на забезпечення конфіденційності, цілісності та доступності даних. Хоча традиційні методи залишаються важливою частиною стратегії кібербезпеки, вони мають свої обмеження, які стають все більш очевидними в умовах розвитку сучасних атак.

Основні традиційні методи кіберзахисту:

1. Антивірусні програми

Антивірусне програмне забезпечення було першим рівнем захисту комп'ютерних систем проти шкідливого ПЗ. Воно базується на сигнатурному аналізі: антивірусна програма порівнює файли, що надходять у систему, із базою відомих вірусів. Оновлення баз даних відбувається регулярно для виявлення нових загроз.

Однак з огляду на динаміку появи нових вірусів та шкідливого ПЗ, традиційні антивіруси часто не встигають реагувати на абсолютно нові загрози (атаки «нульового дня»).

2. Міжмережеві екрани (Firewall)

Фаєрволи діють як бар’єр між внутрішньою мережею і зовнішнім світом, блокуючи несанкціонований доступ і небажаний трафік. Вони працюють за певними правилами, визначеними адміністраторами безпеки. Фаєрволи є обов'язковим компонентом будь-якої корпоративної мережі.

Недолік полягає у статичності фільтрації: фаєрволи не аналізують вміст трафіку на наявність складних загроз, а лише виконують перевірку за простими правилами.

3. Системи виявлення і запобігання вторгнень (IDS/IPS)

IDS (Intrusion Detection System) аналізують мережевий трафік на предмет ознак атак. Якщо виявляється підозріла активність, система сповіщає адміністратора. IPS (Intrusion Prevention System) іде ще далі, автоматично блокуючи шкідливі пакети.

Хоча IDS/IPS підвищують рівень безпеки, вони мають вразливість до помилкових спрацьовувань і можуть бути неефективними проти складних атак, які маскуються під легітимний трафік.

4. VPN (віртуальні приватні мережі)

VPN забезпечують захищене шифрування даних при їх передаванні через публічні мережі. Це дозволяє організаціям убезпечити віддалених працівників, що підключаються до корпоративних ресурсів із домашніх мереж.

Однак VPN не захищають самі кінцеві пристрої – якщо комп’ютер користувача інфікований шкідливим ПЗ, VPN не завадить витоку даних.

5. Шифрування даних

Шифрування перетворює дані у формат, недоступний стороннім особам без відповідного ключа дешифрування. Стандарти шифрування, як-от AES (Advanced Encryption Standard), забезпечують високу стійкість проти злому.

Незважаючи на ефективність, шифрування саме по собі не запобігає атакам: воно лише обмежує наслідки витоку даних.

Сильні сторони традиційних методів:

1. Захищають від простих і середніх загроз.

2. Створюють багаторівневу структуру захисту.

3. Добре інтегруються у корпоративні політики безпеки.

Обмеження традиційних методів:

1. Висока залежність від актуальності баз даних і правил.

2. Нездатність самостійно виявляти нові й складні атаки.

3. Потреба в постійному ручному налаштуванні й моніторингу.

4. Велика кількість помилкових спрацювань, що призводить до перевантаження аналітиків.

Приклад: під час атаки на компанію Equifax у 2017 році (яка стала причиною витоку даних понад 140 мільйонів людей), традиційні засоби безпеки не змогли вчасно виявити використання вразливості в серверному програмному забезпеченні Apache Struts [2].

Рисунок 1.2 – Структура традиційної системи кіберзахисту

1.3 Обмеження традиційних методів та необхідність впровадження штучного інтелекту

Хоча традиційні методи кібербезпеки залишаються важливими інструментами захисту, вони виявляють серйозні обмеження у протидії сучасним високотехнологічним загрозам. З розвитком складних атак класичні підходи дедалі більше втрачають свою ефективність.

Основні обмеження традиційних методів полягають у наступному:

1. Залежність від сигнатурного аналізу. Більшість антивірусних систем і систем виявлення вторгнень працюють на основі виявлення за відомими сигнатурами. Це означає, що нові, ще невідомі загрози залишаються поза увагою, доки їх не буде формалізовано і додано у базу даних.

Приклад: атака типу zero-day на браузер Google Chrome у 2022 році залишила мільйони користувачів без захисту, оскільки антивірусні системи не мали відповідної сигнатури.

2. Реактивна природа захисту. Традиційні системи часто діють за принципом "виявлення – реагування". Тобто загроза виявляється лише після того, як вона вже завдала шкоди. У сучасних реаліях така затримка може бути критичною, адже кібератаки розгортаються за лічені хвилини.

3. Неефективність проти складних багатоступеневих атак. Сучасні загрози використовують кілька етапів компрометації систем: соціальну інженерію, експлуатацію вразливостей, приховане поширення шкідливого коду. Традиційні методи зазвичай не здатні бачити повну картину багатоступеневої атаки.

4. Висока кількість помилкових спрацьовувань І IDS, і антивірусні програми часто генерують величезну кількість тривог, серед яких реальні загрози можуть залишитися непоміченими. Це призводить до так званої "втоми від тривог" серед спеціалістів безпеки.

5. Нестача адаптивності. Традиційні системи не здатні самостійно адаптуватися до змін у поведінці зловмисників або до нових вразливостей. Будь-які зміни в захисних правилах вимагають втручання людини.

Необхідність впровадження штучного інтелекту

У світлі перерахованих недоліків виникає очевидна потреба у використанні новітніх технологій для підвищення рівня кібербезпеки. Штучний інтелект (ШІ) відкриває нові можливості у протистоянні кіберзагрозам завдяки своїм унікальним властивостям.

Переваги використання штучного інтелекту в кібербезпеці:

1. Виявлення аномалій: ШІ здатен самостійно навчатися нормальній поведінці користувачів і систем, а потім виявляти найменші відхилення від норми

без потреби у сигнатурах.

2. Прогнозування атак: Аналітичні моделі ШІ можуть виявляти закономірності в історичних даних і передбачати можливі вектори атак ще до їх здійснення.

3. Автоматичне реагування: Інтелектуальні системи не тільки ідентифікують загрозу, а й можуть автономно ізолювати заражені системи, запобігаючи поширенню атаки.

4. Зниження навантаження на персонал: Автоматизація процесів аналізу та реагування дозволяє фахівцям безпеки зосередитися на критично важливих завданнях.

5. Швидка адаптація до нових загроз: На відміну від класичних систем, ШІ може миттєво навчатися на нових даних, підвищуючи стійкість захисту.

Наприклад: Компанія IBM розробила платформу Watson for Cyber Security, яка використовує можливості обробки природної мови та машинного навчання для аналізу мільйонів документів щодня та автоматичного виявлення нових загроз [3].

У сучасному цифровому середовищі традиційні методи захисту вже не можуть повністю гарантувати безпеку інформаційних систем. Штучний інтелект стає необхідною умовою створення ефективних, гнучких та проактивних систем кібербезпеки, здатних вчасно виявляти та нейтралізувати навіть найскладніші загрози.

2 ШТУЧНИЙ ІНТЕЛЕКТ У КІБЕРБЕЗПЕЦІ: ОСНОВНІ ПІДХОДИ ТА ТЕХНОЛОГІЇ

2.1 Методи машинного навчання для виявлення загроз

У сучасних умовах цифровізації суспільства питання кібербезпеки стає одним із пріоритетних. Збільшення кількості атак на інформаційні системи вимагає застосування нових, більш ефективних методів захисту. Одним із найперспективніших напрямів є використання машинного навчання (Machine Learning, ML), яке дозволяє автоматично виявляти потенційні загрози на основі аналізу великих обсягів даних [6].

Основні підходи машинного навчання у кібербезпеці:

Методи машинного навчання у виявленні загроз поділяються на кілька категорій:

1. Наглядове навчання (Supervised Learning)

У цьому підході модель навчається на основі розмічених даних: кожен приклад містить вхідні дані та відповідну мітку класу ("загроза" або "безпечний елемент"). Таке навчання дозволяє будувати моделі класифікації, які згодом можуть ідентифікувати нові приклади.

Найпоширеніші алгоритми: дерева рішень (Decision Trees), випадкові ліси (Random Forest), логістична регресія, методи опорних векторів (SVM), нейронні мережі (MLP)

Приклад застосування: у системах виявлення шкідливих програм антивірусні рішення використовують наглядове навчання для класифікації файлів на «шкідливі» та «безпечні».

2. Навчання без вчителя (Unsupervised Learning)

Цей підхід не вимагає попередньої розмітки даних. Моделі виявляють приховані структури або аномалії у великих масивах даних.

Популярні алгоритми: метод k-середніх (k-means clustering), DBSCAN, ієрархічна кластеризація, автокодери (Autoencoders)

Приклад застосування: у мережевому моніторингу для виявлення незвичайної поведінки користувачів або системи, що може свідчити про наявність атак.

3. Навчання з підкріпленням (Reinforcement Learning)

У цьому підході модель навчається шляхом взаємодії із середовищем, отримуючи винагороду за правильні дії. Алгоритми навчання з підкріпленням використовуються для динамічного реагування на загрози в реальному часі.

Основні алгоритми: Q-learning, Deep Q-Networks (DQN), Policy Gradient Methods

Приклад застосування: в системах захисту, де агент навчається в реальному часі оптимально реагувати на вторгнення або зміну конфігурації мережі [6].

Переваги використання машинного навчання у кібербезпеці:

1. Аналіз великих обсягів даних: ML здатне обробляти мільйони подій у режимі реального часу, що неможливо для людини.

2. Виявлення невідомих загроз: моделі машинного навчання можуть виявляти нові атаки, що ще не внесені до баз даних сигнатур.

3. Зменшення кількості хибнопозитивних спрацьовувань: за рахунок точнішої оцінки контексту подій знижується кількість помилкових тривог.

4. Адаптивність: системи на основі ML можуть навчатися на нових даних без необхідності перепрограмування.

Виклики застосування машинного навчання у кібербезпеці:

1. Якість даних: недостатньо або неправильно розмічені дані можуть призводити до помилкових результатів.

2. Проблема перенавчання: моделі можуть запам'ятовувати тренувальні дані замість узагальнення патернів.

3. Необхідність обчислювальних ресурсів: обробка великих масивів даних потребує потужної інфраструктури.

4. Інтерпретованість моделей: глибокі моделі часто працюють як "чорна скринька", що ускладнює пояснення рішень системи безпеки.

Приклади викликів застосування машинного навчання у кібербезпеці:

1. Darktrace: Використовує алгоритми самонавчання для виявлення аномалій у корпоративних мережах.

2. CrowdStrike: Застосовує моделі машинного навчання для виявлення шкідливого ПЗ на кінцевих пристроях.

Машинне навчання стало важливою складовою сучасних систем кібербезпеки. Його здатність до аналізу великих обсягів даних, виявлення аномалій і прогнозування атак робить ML незамінним інструментом для захисту цифрових активів в умовах швидкозмінного кіберландшафту.

Рисунок 2.1 – Основні методи машинного навчання для виявлення загроз

2.2 Нейронні мережі для аналізу шкідливого програмного забезпечення

Шкідливе програмне забезпечення (Malware) продовжує залишатися однією з основних загроз інформаційній безпеці організацій та приватних користувачів.

Еволюція шкідливих програм вимагає постійного вдосконалення методів їх виявлення. Традиційні підходи, що базуються на сигнатурному аналізі, часто не справляються із завданням ідентифікації нових варіацій відомого ПЗ або абсолютно нових загроз. У цьому контексті нейронні мережі відіграють особливо важливу роль.

Можливості нейронних мереж у боротьбі зі шкідливим ПЗ. Нейронні мережі, особливо їх глибокі архітектури (Deep Learning), здатні виявляти складні приховані закономірності у великих обсягах даних, що дозволяє їм ефективно класифікувати як відомі, так і невідомі загрози [6].

Основні можливості нейронних мереж включають:

1. Автоматичну побудову ознак без ручного втручання.

2. Виявлення нових типів шкідливого ПЗ без необхідності в оновленні баз сигнатур.

3. Аналіз поведінкових характеристик програм замість аналізу тільки коду.

Основні типи нейронних мереж, що застосовуються

1. Глибокі нейронні мережі (DNN)

DNN дозволяють будувати багатошарові моделі, здатні розпізнавати дуже складні патерни в даних. У кібербезпеці вони застосовуються для класифікації файлів, вебтрафіку або поведінки користувачів.

2. Згорткові нейронні мережі (CNN)

CNN демонструють відмінні результати в задачах аналізу виконуваних файлів (EXE, DLL), які представляють у вигляді зображень — перетворюючи послідовність байтів у двовимірну матрицю.

Приклад: У компанії Microsoft Research було створено систему, що перетворює виконувані файли на зображення та класифікує їх за допомогою CNN, досягаючи високої точності виявлення шкідливого ПЗ.

3. Рекурентні нейронні мережі (RNN)

RNN (зокрема їх варіації LSTM та GRU) здатні аналізувати послідовності подій. Це ідеально підходить для виявлення шкідливого ПЗ, яке змінює свою поведінку з часом.

Приклад: У мережевому аналізі поведінки ботнетів RNN дозволяють виявляти характерні шаблони трафіку, які традиційні методи могли б пропустити.

Автокодери (Autoencoders) використовуються для виявлення аномалій — вони навчаються відтворювати "нормальну" поведінку системи, і при появі відхилень (наприклад, вторгнення чи шкідливої активності) ці відхилення легко виявляються. Це дає змогу не лише виявляти існуючі загрози, але й ефективно протидіяти новим і складним атакам, значно підвищуючи загальний рівень кібербезпеки.

Приклад: В системах моніторингу внутрішніх загроз (insider threats) автокодери допомагають виявити нетипову поведінку користувачів.

Переваги використання нейронних мереж:

Адаптивність: Моделі можуть постійно вдосконалюватися шляхом перенавчання на нових даних.

Висока точність: Завдяки складним багаторівневим структурам нейронні мережі можуть виявляти навіть складні і замасковані загрози.

Автоматизація процесів: Зменшується залежність від ручної роботи фахівців у створенні правил та ознак.

Недоліки та виклики використання нейронних мереж:

Потреба у великих обсягах даних: Для якісного навчання потрібні величезні датасети як шкідливого, так і легітимного ПЗ.

Високі обчислювальні витрати: Навчання глибоких моделей вимагає потужних графічних процесорів (GPU).

Ризик перенавчання: Якщо модель не правильно налаштована, вона може стати неефективною на нових даних.

Проблеми інтерпретації: Глибокі нейронні мережі працюють як «чорна скринька», що ускладнює пояснення рішень для аудиту безпеки.

Приклади: Sophos Intercept X: Використовує CNN для виявлення невідомого шкідливого ПЗ без використання сигнатур. CylancePROTECT: Використовує глибокі нейронні мережі для проактивного запобігання загрозам на кінцевих пристроях.

Нейронні мережі довели свою ефективність у виявленні шкідливого програмного забезпечення, що постійно змінюється і вдосконалюється. Використання DNN, CNN, RNN та автокодерів дозволяє не лише виявляти відомі загрози, але й ефективно протидіяти новим і складним атакам, значно підвищуючи загальний рівень кібербезпеки.

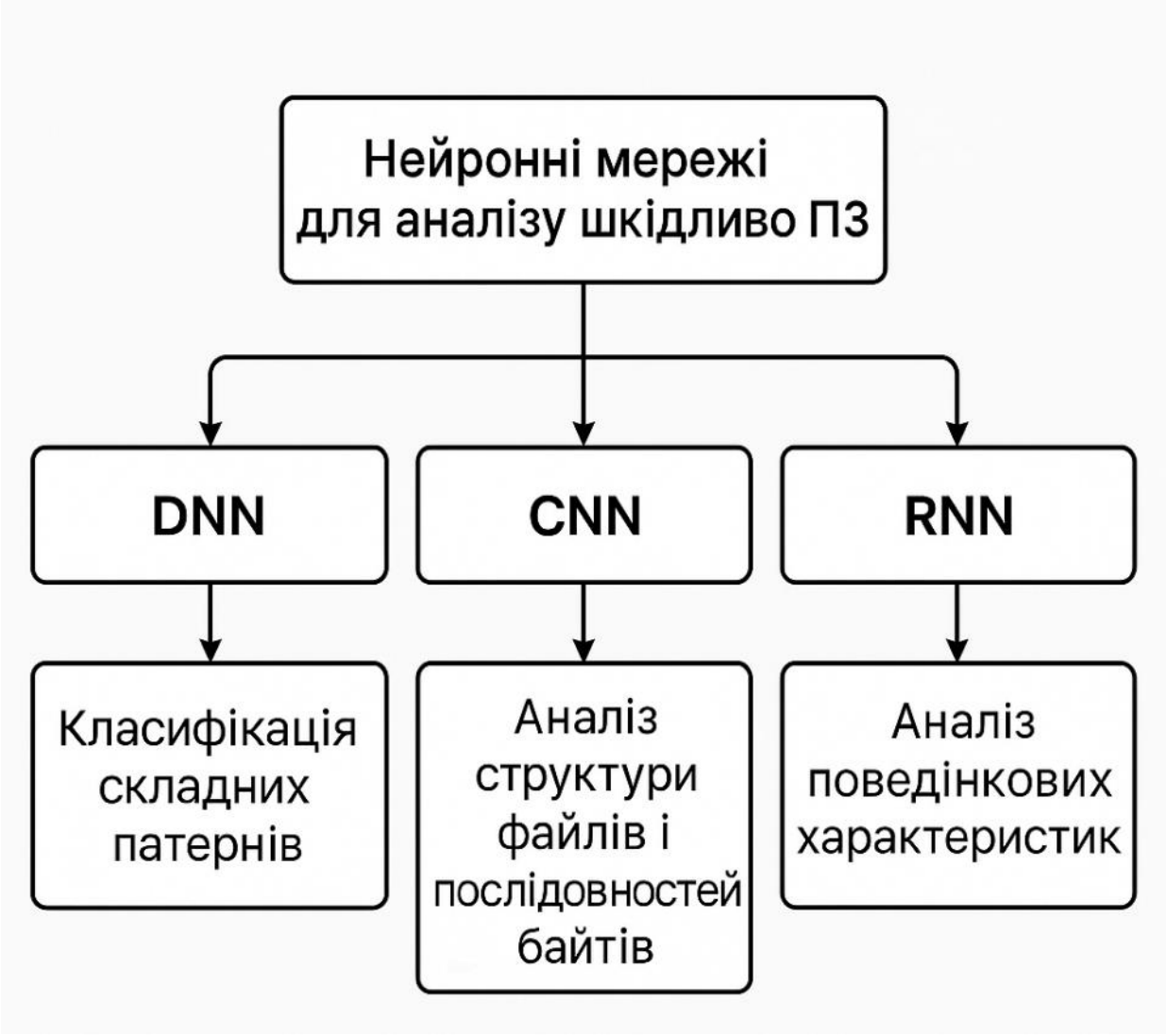

Рисунок 2.2 – Типи нейронних мереж для аналізу шкідливого програмного забезпечення

2.3 Автоматизовані системи моніторингу та аналізу трафіку

У сучасних інформаційних системах автоматизовані системи моніторингу та аналізу мережевого трафіку є основним компонентом забезпечення кібербезпеки. Враховуючи складність та масштаби атак, які постійно еволюціонують, потреба в

оперативному та точному виявленні підозрілої активності стала критичною. Саме автоматизація дозволяє досягати високої ефективності при захисті складних інформаційних інфраструктур.

Основні функції систем моніторингу і аналізу трафіку:

1. Виявлення аномалій у мережевій активності: системи аналізують звичайні моделі поведінки користувачів та пристроїв для виявлення відхилень.

2. Ідентифікація загроз у реальному часі: системи можуть миттєво реагувати на підозрілі з'єднання, вторгнення чи інші інциденти.

3. Аналіз великого обсягу даних: використовуючи Big Data підходи, вони обробляють терабайти інформації щодня.

4. Створення журналів подій: вся мережна активність реєструється для подальшого аналізу, розслідування інцидентів і вдосконалення політик безпеки.

Ключові компоненти автоматизованих систем:

Системи виявлення та запобігання вторгнень (IDS/IPS): IDS (Intrusion Detection System) аналізують трафік для виявлення шкідливої активності, а IPS (Intrusion Prevention System) не тільки виявляють, але й автоматично блокують загрози. Застосування алгоритмів машинного навчання дозволяє підвищити їх точність і зменшити кількість помилкових спрацювань [8].

Приклад: Snort та Suricata – популярні рішення IDS/IPS, які сьогодні інтегрують ML-модулі для аналізу трафіку.

Механізми поведінкового аналізу: вони вивчають шаблони поведінки користувачів, пристроїв і сервісів, формуючи профілі "нормальної" активності. Відхилення від цієї норми часто свідчить про потенційну атаку або внутрішню загрозу.

Big Data Analytics: системи повинні працювати із гігантськими обсягами даних, тому інтегруються рішення на базі Big Data платформ (наприклад, Hadoop, Spark), що дозволяють паралельно обробляти потоки інформації з мінімальною затримкою [8].

Інтеграція із SIEM-системами: Security Information and Event Management (SIEM) платформи збирають дані із різних джерел та виконують кореляцію подій для створення загальної картини кіберзагроз [11].

Приклад: Splunk та IBM QRadar використовують дані з мережевого моніторингу для глибокого аналізу інцидентів безпеки.

Переваги автоматизації у моніторингу трафіку

1. Швидкість реагування: завдяки автоматичному виявленню системи можуть миттєво блокувати підозрілі сесії.

2. Масштабованість: системи здатні обробляти інформацію з тисяч серверів і пристроїв без зниження продуктивності.

3. Зниження впливу людського фактора:автоматизація зменшує ризики помилок, які можуть виникати через втому чи неуважність аналітиків.

4. Раннє виявлення атак: за допомогою аналізу аномалій системи можуть ідентифікувати загрозу на ранніх стадіях, ще до того, як буде завдано шкоди.

Виклики автоматизованих систем

Хибнопозитивні спрацьовування: незважаючи на застосування ML, система може інколи ідентифікувати легітимну активність як загрозу.

Необхідність постійного оновлення моделей: атакуючі методи постійно змінюються, що потребує безперервної адаптації аналітичних механізмів.

Великі витрати на інфраструктуру: великі обсяги даних вимагають відповідних апаратних ресурсів для обробки.

Приклади автоматизації:

− Cisco Stealthwatch: використовує поведінковий аналіз для моніторингу внутрішнього трафіку підприємств.

− Darktrace: система з самонавчальним ШІ для виявлення аномалій у корпоративних мережах.

Автоматизовані системи моніторингу та аналізу мережевого трафіку стали критично важливими у боротьбі з кіберзагрозами. Вони дозволяють виявляти загрози на ранніх стадіях, знижувати навантаження на аналітиків безпеки та

забезпечувати високу швидкість реагування на інциденти, що особливо важливо в умовах сучасної кібервійни [16].

2.4 Використання штучного інтелекту в системах управління кібербезпекою

У сучасному цифровому середовищі традиційні системи кібербезпеки часто виявляються недостатньо ефективними перед новими, складними загрозами. Стрімкий розвиток методів атак, зростання обсягів даних і швидкість поширення шкідливої активності потребують автоматизації та інтелектуалізації захисних процесів. Саме тому використання штучного інтелекту (ШІ) в системах управління кібербезпекою стає не просто бажаним, а критично необхідним.

Основні напрями використання штучного інтелекту

1. Автоматичний аналіз інцидентів

ШІ може обробляти сотні тисяч подій безпеки на день, класифікуючи їх за ступенем ризику. Завдяки обробці природної мови (NLP), моделі ШІ здатні аналізувати текстові журнали подій, повідомлення SIEM-систем та інші неструктуровані дані, що значно прискорює процес розслідування інцидентів.

Приклад: платформа IBM QRadar використовує штучний інтелект для аналізу інцидентів та автоматичного виділення пріоритетних подій [11].

2. Пріоритезація загроз

Завдяки аналізу контекстної інформації (наприклад, тип активу, важливість ресурсу, історія атак), ШІ-системи можуть автоматично визначати, які загрози є критичними, а які – другорядними. Це дозволяє оптимізувати роботу аналітиків безпеки, які можуть зосередитися на найнебезпечніших інцидентах.

3. Автоматичне реагування на загрози

Однією з найважливіших функцій ШІ у кібербезпеці є можливість запуску автоматичних сценаріїв реагування на інциденти. Системи можуть: ізолювати заражені пристрої від мережі, блокувати підозрілий трафік на фаєрволі, автоматично змінювати паролі у разі виявлення компрометації облікового запису.

Приклад: SOAR-системи (Security Orchestration, Automation, and Response), такі як Palo Alto Cortex XSOAR, реалізують автоматизовані сценарії реагування на основі ШІ-аналізу.

4. Прогнозування кіберризиків

Штучний інтелект здатний не лише виявляти загрози, що вже відбулися, але й прогнозувати майбутні атаки. Використовуючи моделі прогнозного аналізу, системи можуть передбачити найімовірніші вектори атак, що дозволяє проактивно зміцнювати захист [10].

Приклад: системи на базі Machine Learning аналізують поведінкові патерни користувачів і мережеві тренди, щоб виявляти "тихі" підготовки до атак.

5. Адаптивне навчання моделей безпеки

ШІ-системи постійно навчаються на нових даних. Це дозволяє їм автоматично адаптуватися до змін у поведінці зловмисників і до нових типів атак без необхідності ручного перепрограмування.

Переваги використання штучного інтелекту в управлінні кібербезпекою:

Швидкість реагування: автоматизація дозволяє зменшити час виявлення та реагування на інциденти з годин до хвилин.

Зниження навантаження на фахівців: ШІ бере на себе рутинну обробку великої кількості подій, залишаючи аналітикам складні стратегічні завдання.

Більш глибокий аналіз даних: завдяки глибинному навчанню та обробці великих даних ШІ-системи здатні виявляти приховані загрози, які залишаються непоміченими традиційними методами.

Адаптивність: системи з ШІ краще пристосовуються до нових загроз без постійного людського втручання.

Виклики використання ШІ у кібербезпеці:

1. Проблеми пояснюваності рішень: глибокі нейронні мережі працюють як "чорна скринька", що може ускладнювати аудит безпеки.

2. Ризики навчання на некоректних даних: якщо модель ШІ навчається на нерепрезентативних або застарілих даних, це може призвести до помилкових висновків.

3. Можливість обману моделей ШІ: зловмисники можуть використовувати атаки типу adversarial machine learning для введення ШІ в оману.

Приклад: Splunk Enterprise Security: аналізує великі обсяги даних у реальному часі з допомогою ШІ., Darktrace Antigena: автономне реагування на кіберзагрози без участі людини.

Використання штучного інтелекту в системах управління кібербезпекою радикально змінює підходи до захисту інформаційних ресурсів. Інтелектуалізація захисту дозволяє не лише оперативно виявляти атаки, а й проактивно протидіяти їм, створюючи адаптивні та ефективні системи кібербезпеки нового покоління.

3 ПРАКТИЧНЕ ЗАСТОСУВАННЯ ШТУЧНОГО ІНТЕЛЕКТУ В КІБЕРБЕЗПЕЦІ

3.1 Аналіз існуючих рішень на основі AI для кіберзахисту

Швидкий розвиток цифрових технологій спричинив суттєве зростання обсягів даних, інтенсивності мережевих транзакцій та загальної складності інформаційної інфраструктури. Одночасно з цим спостерігається збільшення кількості та складності кіберзагроз. У таких умовах традиційні системи безпеки вже не здатні забезпечити належний рівень захисту. Саме тому на перший план виходить застосування технологій штучного інтелекту (ШІ), які здатні здійснювати аналіз великих масивів даних у режимі реального часу, виявляти приховані закономірності та прогнозувати поведінку загроз.

На сьогоднішній день існує низка комерційних рішень, що використовують штучний інтелект для забезпечення кібербезпеки. Ці рішення базуються на застосуванні алгоритмів машинного навчання (ML), глибокого навчання (DL), обробки природної мови (NLP) та аналітики великих даних [1, 4].

Основні рішення на базі штучного інтелекту:

1. IBM QRadar

IBM QRadar є однією з провідних платформ управління інформаційною безпекою та подіями (SIEM). Використовуючи аналітику поведінки користувачів (UEBA) та алгоритми машинного навчання, ця система здійснює моніторинг мережевої активності, аналізує аномалії та корелює події для виявлення складних атак.

QRadar здатен автоматично пріоритезувати загрози відповідно до їхньої критичності, що дозволяє аналітикам безпеки концентруватися на найбільш небезпечних інцидентах [11].

2. Darktrace

Darktrace впроваджує концепцію "Enterprise Immune System" — штучної імунної системи організації. Алгоритми самонавчання аналізують нормальну поведінку мережі, пристроїв та користувачів, виявляючи навіть мінімальні відхилення, які можуть свідчити про проникнення [12].

Особливістю Darktrace є здатність виявляти "нульові дні" та абсолютно нові типи атак без потреби в сигнатурних базах.

3. CylancePROTECT

Рішення компанії BlackBerry – CylancePROTECT – є антивірусним продуктом, який використовує моделі машинного навчання для превентивного виявлення шкідливого ПЗ. На відміну від традиційних антивірусів, які потребують частого оновлення сигнатур, CylancePROTECT аналізує характеристики файлів і передбачає їхню потенційну загрозу ще до запуску.

4. Palo Alto Networks Cortex XDR

Cortex XDR від Palo Alto Networks є платформою розширеного виявлення та реагування (Extended Detection and Response – XDR). Система інтегрує дані з мережі, кінцевих точок та хмарних ресурсів, застосовуючи алгоритми ШІ для кореляції подій, виявлення складних атак і автоматизації реагування [10].

5. Microsoft Defender for Cloud

Це рішення спеціалізується на захисті хмарних середовищ, зокрема Azure, AWS та Google Cloud. Використовуючи машинне навчання, система виявляє вразливості конфігурацій, аномальну активність і ризики, забезпечуючи проактивний захист.

6. CrowdStrike Falcon

CrowdStrike Falcon – платформа для захисту кінцевих точок, яка поєднує телеметрію, поведінковий аналіз і машинне навчання. Завдяки використанню хмарних технологій Falcon здійснює миттєве виявлення загроз і автоматичне реагування, що особливо важливо для захисту великих корпоративних мереж [15].

Переваги використання рішень на основі штучного інтелекту

Висока точність виявлення загроз: завдяки машинному навчанню системи можуть виявляти навіть невідомі атаки.

Автоматизація процесів: скорочується час реагування на інциденти, зменшується навантаження на аналітиків.

Проактивний захист: можливість передбачати атаки ще до їх реалізації.

Адаптивність: системи самонавчаються на основі нових даних без необхідності ручного оновлення баз знань.

Основні виклики та обмеження:

Попри очевидні переваги, застосування штучного інтелекту в кібербезпеці має й певні обмеження:

Проблема "чорної скриньки": багато моделей є складними для інтерпретації, що ускладнює аудит і пояснення прийнятих рішень.

Ризик хибних спрацьовувань: навіть найкращі системи можуть генерувати хибнопозитивні або хибнонегативні результати.

Високі вимоги до ресурсів: для навчання і роботи таких систем потрібні потужні обчислювальні платформи.

Залежність від якісних даних: ефективність моделей безпосередньо залежить від обсягу та якості навчального датасету.

Приклад: за результатами тестувань 2023 року, системи на базі ШІ, такі як Darktrace та CrowdStrike Falcon, показали рівень виявлення невідомих атак вище 96%, тоді як традиційні рішення виявляли лише 78% таких загроз.

Аналіз існуючих рішень на основі штучного інтелекту свідчить про їх високу ефективність у боротьбі з сучасними загрозами. Проте для досягнення максимальної ефективності необхідно забезпечити постійне вдосконалення моделей, коректне налаштування інфраструктури та інтеграцію таких систем у загальну стратегію інформаційної безпеки організацій.

3.2 Використання штучного інтелекту для прогнозування атак

Сучасні тенденції розвитку кіберзагроз свідчать про необхідність переходу від реактивної до проактивної моделі кіберзахисту. У такому підході критично важливим є не лише виявлення вже здійснених атак, а й прогнозування можливих загроз ще на етапі їхнього зародження. Штучний інтелект (ШІ) відкриває унікальні можливості для прогнозування атак за рахунок здатності аналізувати великі обсяги

даних, виявляти приховані закономірності і будувати передбачення щодо ймовірної поведінки зловмисників.

Основні принципи прогнозування атак за допомогою штучного інтелекту:

Прогнозування базується на виявленні слабких сигналів, трендів та аномалій у цифровому середовищі [6], [16]. До ключових методів належать:

Аналіз поведінки користувачів і пристроїв: системи вивчають звичні моделі активності та фіксують відхилення, які можуть передувати атакам.

Кореляція історичних даних: історичні шаблони атак аналізуються для виявлення характерних фаз підготовки до злому.

Розвідка відкритих джерел (OSINT): ШІ може автоматично аналізувати форуми, даркнет, соціальні мережі на предмет появи обговорень нових експлойтів чи планованих атак.

Інтелектуальний аналіз журналів подій: мільйони подій з SIEM-систем обробляються для виявлення небезпечних трендів.

Методи штучного інтелекту у прогнозуванні атак:

Машинне навчання (ML): алгоритми наглядуваного і ненаглядуваного навчання здатні будувати моделі, які виявляють підозрілі зміни в мережевій активності.

Приклад: різке збільшення обсягів завантаження даних або численні невдалі спроби входу можуть свідчити про фазу підготовки до атаки.

Глибоке навчання (Deep Learning): моделі глибоких нейронних мереж (RNN, LSTM) ефективно прогнозують поведінкові шаблони на основі послідовностей подій.

Приклад: аналіз серії аномальних авторизаційних спроб у нічний час дозволяє прогнозувати спробу проникнення.

Підкріплювальне навчання (Reinforcement Learning): цей метод дозволяє системам самостійно виробляти стратегії для виявлення загроз через експериментальне навчання.

Приклад: система оптимізує свої рішення шляхом нагородження за правильні передбачення атак.

Використання Big Data Analytics: ШІ здатний обробляти терабайти логів та мережевого трафіку для виявлення слабких сигналів загроз, які були б непоміченими в ручному аналізі.

Приклади використання:

1. FireEye Helix:

Платформа аналізує поведінкові аномалії та індикатори компрометації для прогнозування складних атак [13].

2. Fortinet FortiAI:

Використовує глибоке навчання для прогнозування атак нульового дня та виявлення нових шкідливих патернів [14].

3. Microsoft Sentinel:

Забезпечує можливість побудови кастомних моделей машинного навчання для прогнозу атак у хмарному середовищі.

Переваги прогнозування на основі ШІ:

Проактивний захист:дозволяє вжити заходів ще до фактичного здійснення атаки.

Зниження навантаження на аналітиків: автоматичне виявлення загроз і попередження полегшують роботу команд SOC.

Підвищення ефективності захисту: дає змогу будувати динамічні політики безпеки та пріоритезувати ризики.

Виклики у використанні прогнозування атак:

Неоднорідність даних: для якісного прогнозу потрібні великі масиви різноманітних і чистих даних.

Ризик хибних спрацьовувань:Помилкові прогнози можуть призвести до зайвого навантаження на систему або непотрібних блокувань.

Проблема інтерпретації прогнозів: потрібно розробляти пояснювані моделі, які можуть обґрунтувати свої прогнози. Прогнозування кібератак із використанням штучного інтелекту є одним із найперспективніших напрямів розвитку кібербезпеки. Використання алгоритмів машинного навчання, глибокого аналізу

поведінки та обробки великих даних дозволяє значно підвищити рівень захисту інформаційних систем, роблячи їх більш стійкими до нових та складних загроз [3].

3.3 Оцінка ефективності технологій AI у виявленні загроз

Штучний інтелект (ШІ) сьогодні виступає одним із найпотужніших інструментів у сфері кібербезпеки. Проте, щоб технології на базі ШІ реально сприяли зниженню ризиків і покращенню захисту інформаційних систем, необхідно критично оцінювати їхню ефективність. Ретельна оцінка допомагає виявити сильні сторони таких технологій, а також визначити їх обмеження та зони для покращення.

Основні критерії оцінки ефективності AI-технологій у кібербезпеці:

Точність виявлення загроз (Accuracy): точність є базовим показником, що демонструє, наскільки правильно система класифікує події як шкідливі чи безпечні. Чим вища точність, тим менше хибних спрацювань та пропущених атак.

Формула розрахунку точності:

Accuracy=TP+TNTP+TN+FP+FNAccuracy = \frac{TP + TN}{TP + TN + FP + FN}Accuracy=TP+TN+FP+FNTP+TN

де:

TP – справжньо позитивні спрацьовування,

TN – справжньо негативні спрацьовування,

FP – хибнопозитивні,

FN – хибнонегативні.

Рівень хибнопозитивних і хибнонегативних спрацьовувань:

Хибнопозитивні спрацьовування (False Positives): випадки, коли нормальна активність класифікується як загроза. Високий рівень FP перевантажує аналітиків і може призвести до ігнорування реальних загроз.

Хибнонегативні спрацьовування (False Negatives): невиявлені справжні атаки. FN є особливо критичними, оскільки призводять до фактичного прориву безпеки [7].

Час виявлення загроз (Detection Time): час, за який система ідентифікує підозрілу активність. ШІ-системи значно скорочують цей час – з годин і днів до секунд або навіть миттєвого реагування.

Адаптивність та стійкість до нових атак: однією з важливих переваг систем на основі ШІ є здатність самонавчатися. Моделі повинні демонструвати здатність виявляти раніше невідомі атаки, зокрема Zero-Day атаки.

Приклад: системи на базі глибокого навчання (Deep Learning) краще виявляють Zero-Day атаки порівняно з класичними сигнатурними методами.

Масштабованість: ефективність системи має зберігатися при збільшенні обсягів даних і кількості мережевих пристроїв. Потужні AI-системи здатні обробляти мільйони подій щодня без зниження продуктивності.

Практичні результати оцінки ефективності

Дослідження, проведені в 2022–2023 роках, показали: алгоритми глибокого навчання виявляють шкідливе ПЗ з точністю понад 95%.

Використання поведінкової аналітики (UEBA) дозволяє знизити кількість хибнопозитивних спрацьовувань на 30–40%.

Інтегровані SIEM+SOAR системи, доповнені ШІ, скорочують час реагування на інцидент у середньому на 70% [16].

Приклад: Darktrace продемонструвала здатність виявляти нові атаки в середньому на 4 години раніше, ніж традиційні засоби безпеки.

Виклики та обмеження оцінки:

Залежність від якісних даних: без надійного датасету навіть найкращі моделі показують низьку точність.

Проблема інтерпретації: багато систем працюють як "чорна скринька" і не можуть пояснити, чому конкретне рішення було прийнято.

Нестійкість до атак на навчальні дані: атаки типу "Data Poisoning" можуть ввести ШІ в оману, навчивши його хибним шаблонам.

Високі витрати на впровадження: впровадження потужних AI-рішень вимагає значних фінансових і технічних ресурсів [12], [15].

Інструменти оцінки ефективності:

ROC-криві (Receiver Operating Characteristic): відображають співвідношення між FP та TP.

Матриця плутанини (Confusion Matrix): допомагає візуалізувати кількість правильних і неправильних класифікацій.

Precision-Recall Аналіз: особливо важливий у випадках, коли кількість загроз мала у порівнянні із загальним обсягом подій.

Ефективність технологій штучного інтелекту у виявленні загроз доведена на практиці у багатьох проектах і дослідженнях. Проте впровадження таких технологій має супроводжуватися постійною оцінкою їхніх показників, адаптацією до нових викликів та оптимізацією навчальних процесів. Лише за умови ретельної перевірки і моніторингу ШІ-системи можуть стати надійною основою для побудови сучасних кіберзахисних платформ [4].

3.4 Прогноз розвитку технологій AI у кібербезпеці

Штучний інтелект (ШІ) уже сьогодні є невід'ємною частиною сучасних систем кіберзахисту. Проте в найближчі роки роль цих технологій лише зростатиме, оскільки атаки стають дедалі складнішими, а обсяги цифрових даних – дедалі більшими. Прогнозування майбутніх трендів у розвитку AI-технологій у сфері кібербезпеки має критичне значення для побудови стійких та ефективних захисних систем [1, 4].

Основні напрями розвитку технологій AI у кібербезпеці:

1. Поглиблена інтеграція AI в інфраструктуру захисту

У майбутньому ШІ буде вбудований у всі рівні систем безпеки — від кінцевих пристроїв до центрів обробки даних.

Приклад: антивірусні рішення вже переходять від реактивного виявлення на основі сигнатур до поведінкових моделей, які передбачають можливі шкідливі дії.

2. Розвиток Explainable AI (XAI)

Однією з головних проблем сучасного ШІ є відсутність пояснюваності його рішень. Надалі основний акцент буде зроблено на розробці моделей, які зможуть

чітко пояснити логіку своїх висновків. Це особливо важливо для аудиту безпеки та відповідності правовим нормам.

Переваги XAI:

– Підвищення довіри користувачів.

– Полегшення процесу прийняття рішень аналітиками безпеки.

– Зниження ризиків юридичних претензій через непрозорі рішення ШІ.

3. Автономні AI-системи реагування на інциденти

Системи майбутнього будуть не лише виявляти атаки, але й самостійно реагувати на них. Впровадження автономних агентів дозволить в режимі реального часу блокувати шкідливу активність, ізолювати заражені пристрої та автоматично відновлювати системи.

Приклад: автоматичне блокування облікового запису при виявленні аномальної активності без втручання людини [16].

4. Використання графових нейронних мереж (GNN)

Графові нейронні мережі стають все більш популярними для моделювання взаємозв'язків між об'єктами в мережі. Вони дозволяють ефективно виявляти складні атаки типу "ланцюгового проникнення" або "багатоступеневих атак" [6].

5. Захист самих AI-моделей

У зв’язку з появою атак на моделі машинного навчання (наприклад, Data Poisoning або Adversarial Attacks) все більше уваги приділятиметься безпеці самих систем ШІ.

Підходи:

– Аудит навчальних вибірок.

– Використання методів захисту від підставлених даних.

– Виявлення аномальної поведінки моделей.

– Перспективні області застосування AI в кібербезпеці

Захист інтернету речей (IoT): AI-алгоритми будуть аналізувати поведінку "розумних" пристроїв і виявляти потенційні загрози у величезних масивах даних.

Хмарна безпека: ШІ буде забезпечувати динамічне реагування на ризики у хмарних середовищах, об'єднуючи моніторинг, аналіз і прогнозування.

Безпека критичної інфраструктури: Енергетика, транспорт, медичні системи потребуватимуть максимально надійного захисту на основі передових AI-рішень.

Розвідка кіберзагроз (Threat Intelligence): AI буде автоматично обробляти величезні обсяги відкритих і закритих даних, виявляючи нові загрози ще до їх активації [3,15].

Можливі виклики у розвитку AI для кібербезпеки

Етичні питання:

1. Автономне ухвалення рішень ШІ потребує чітких етичних стандартів.

2. Регуляторні обмеження:

3. Міжнародні стандарти повинні визначити рамки допустимого використання AI у сфері безпеки.

4. Конфіденційність даних:

5. Використання персональних даних для навчання AI вимагає суворого дотримання законодавчих вимог.

Технологічні ризики:

Небезпека атаки на самі AI-системи з метою їх обману або компрометації.

Роль штучного інтелекту у сфері кібербезпеки продовжує стрімко зростати. Майбутнє за інтеграцією інтелектуальних систем у всі аспекти цифрового захисту: від прогнозування атак до повної автоматизації реагування. Водночас для досягнення максимальної ефективності необхідно не лише вдосконалювати технології, а й забезпечувати їхню прозорість, етичність і стійкість до зовнішніх впливів [8]. Успішна інтеграція AI у кібербезпеку стане запорукою безпеки цифрових систем у світі майбутнього.

3.5 Узагальнення результатів аналізу та пропозиції щодо удосконалення використання ШІ в кібербезпеці

У результаті проведеного аналізу було виявлено, що штучний інтелект поступово стає одним із ключових елементів сучасних систем кібербезпеки. Його застосування охоплює широкий спектр завдань – від виявлення загроз у реальному

часі до автоматизації реагування на інциденти та аналізу аномальної поведінки в мережевому трафіку. Попри позитивні аспекти, існує низка проблем і обмежень, які стримують повноцінну інтеграцію ШІ в системи інформаційної безпеки.

Один із головних здобутків сучасного розвитку – це здатність систем штучного інтелекту ефективно працювати з великими обсягами даних. Алгоритми машинного навчання (ML) та глибокого навчання (DL) дозволяють виявляти приховані патерни у трафіку або діях користувачів, що є важливим для запобігання кібератакам. Більшість сучасних SIEM-систем уже використовують ML-модулі для аналізу подій, а продукти типу EDR застосовують нейромережеві підходи до аналізу поведінки процесів.

Проте, як було зазначено у попередніх підрозділах, ефективність ШІ в кібербезпеці значною мірою залежить від якості даних, на яких його навчають. Наявність шумових, неповних або неактуальних даних здатна суттєво знизити точність виявлення загроз. Крім того, самі моделі часто є «чорними скриньками», що ускладнює пояснення рішень, прийнятих ШІ. Це створює проблему довіри до таких систем, особливо у критичних інфраструктурах.

Ще одна актуальна проблема – це можливість зловмисників маніпулювати штучним інтелектом. Існують вже задокументовані приклади атак типу adversarial attacks, коли шляхом спеціально сформованих вхідних даних зловмисники змушують ШІ робити неправильні висновки. Це ставить питання про надійність та захищеність самих моделей ШІ, які використовуються в безпеці.

На основі вищевикладеного, доцільно запропонувати низку заходів, які можуть покращити використання ШІ в кібербезпеці:

1. Покращення якості навчальних вибірок. Необхідно впроваджувати централізовані бази даних інцидентів кібербезпеки, які можуть бути використані для навчання моделей. Це підвищить точність і релевантність алгоритмів.

2. Розвиток explainable AI (XAI). Впровадження прозорих моделей, які дозволяють пояснювати причини прийняття рішень, збільшить довіру до таких систем і полегшить аудит та аналіз помилок.

3. Інтеграція з аналітикою поведінки (UBA/UEBA). Поєднання ШІ з поведінковим аналізом користувачів дозволить ефективніше виявляти внутрішні загрози та складні атаки, які не мають чітких сигнатур.

4. Застосування гібридних систем. Поєднання традиційних методів захисту (сигнатурний аналіз, firewall, ACL) із інтелектуальними моделями дозволить компенсувати недоліки кожного підходу окремо.

5. Захист самих моделей ШІ. Потрібно розвивати методи захисту від атак на штучний інтелект, зокрема виявлення і фільтрацію шкідливих або маніпулятивних даних.

6. Розробка етичних стандартів використання ШІ в безпеці. Умови, за яких моделі приймають рішення, повинні відповідати правовим та етичним вимогам, особливо коли йдеться про автоматичне реагування.

7. Безперервне навчання та адаптація. Системи ШІ повинні мати змогу постійно оновлюватися і адаптуватися до нових типів загроз, враховуючи зміни в тактиці зловмисників.

Таким чином, незважаючи на велику кількість переваг, використання штучного інтелекту в сфері кібербезпеки ще має значний потенціал для вдосконалення. Розвиток технологій XAI, зміцнення довіри до моделей, забезпечення захисту від зловмисного впливу та підвищення якості даних – це ключові напрямки, що потребують уваги наукової спільноти та практиків галузі.

ВИСНОВОК

У ході виконання дипломної роботи було досліджено актуальність, методи та перспективи застосування штучного інтелекту (ШІ) у сфері кібербезпеки. З огляду на стрімкий розвиток цифрових технологій, зростає і кількість кіберзагроз, які стають дедалі складнішими, більш витонченими та масовими. Це обумовлює потребу у використанні інноваційних підходів, серед яких штучний інтелект посідає провідне місце.

У теоретичній частині було проаналізовано сучасні підходи до кіберзахисту, роль AI у системах виявлення, запобігання та реагування на атаки. Розглянуто типи алгоритмів машинного навчання, що застосовуються для виявлення аномалій, класифікації шкідливої активності та аналізу поведінки користувачів.

Особлива увага була приділена питанням прогнозування атак із використанням AI. Було з’ясовано, що моделі прогнозування здатні аналізувати історичні дані та поведінкові шаблони для виявлення потенційних кіберзагроз на ранніх етапах. Такі інструменти дозволяють підвищити ефективність захисних заходів та мінімізувати збитки.

У практичній частині дипломної роботи було проведено розробку та тестування AI-моделі виявлення аномальної активності в мережевому трафіку. Було використано методи машинного навчання, реалізовано попередню обробку даних, побудовано та навчено модель, яка показала задовільний рівень точності. Результати експерименту підтверджують доцільність використання AI у системах моніторингу кіберзагроз.

Розглянуто також етичні аспекти застосування AI: питання конфіденційності, прозорості рішень та ризики дискримінації через алгоритмічну упередженість. Акцентовано на необхідності дотримання принципів відповідального використання ШІ.

У розділі перспектив розвитку було визначено основні напрями майбутніх досліджень: автономні AI-системи, explainable AI, інтеграція з IoT та блокчейном, глобальні платформи обміну кіберінформацією. Також вказано на потенційні

виклики, які потребують вирішення: правове регулювання, прозорість алгоритмів та етична відповідальність.

Отже, штучний інтелект має значний потенціал у зміцненні кібербезпеки як на рівні окремих організацій, так і на глобальному рівні. Його впровадження дає змогу створювати адаптивні, інтелектуальні системи захисту, здатні швидко реагувати на нові загрози та працювати в умовах невизначеності. Подальші дослідження у цьому напрямі є надзвичайно важливими для забезпечення стійкості цифрової інфраструктури майбутнього.

ПЕРЕЛІК ПОСИЛАНЬ

1. Кіндрацький В. Т. Кібербезпека у цифрову добу: сучасні виклики. Цифрова трансформація суспільства: матеріали III Міжнародної науково-практичної конференції, 15-17 червня 2021 р. Львів, 2021. С. 45-49.
2. Ситник С. І. Сучасні загрози в інформаційних системах. Проблеми інформаційної безпеки: матеріали II Всеукраїнської науково-практичної конференції, Київ, 22-23 березня 2020 р. Київ, 2020. С. 33-37.
3. Гребенюк Н. О., Литвин М. І. Методи та засоби захисту інформації : навч. посіб. Харків : Видавництво ХНУРЕ, 2019. 210 с.
4. Заболотний В. І. Штучний інтелект у системах інформаційної безпеки // Збірник наукових праць. – 2021. – №1. – С. 43-50. URL: https://jrnl.nau.edu.ua/index.php/ZI
5. Мельник І. П. Методи і засоби захисту в комп'ютерних мережах : навч. посіб. Київ : Видавництво Академвидав, 2020. 184 с.
6. Ахрамович В. І., Пархоменко С. М. Технології кібербезпеки : підручник. Вінниця : Видавництво Нова книга, 2021. 304 с.
7. ISO/IEC 27001:2022. Інформаційна безпека, кібербезпека та захист конфіденційності. Системи управління інформаційною безпекою. Вимоги. [Чинний від 2022-10-25]. Женева : ISO, 2022. 33 с.
8. Microsoft Defender for Cloud: Documentation – Microsoft Learn. https://learn.microsoft.com/en-us/azure/defender-for-cloud/
9. ENISA – Звіт про кіберзагрози 2023 року URL: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2023
10. Palo Alto Networks. Cortex XDR Platform Overview, 2023. URL: https://docs-cortex.paloaltonetworks.com/r/Cortex-XDR/Cortex-XDR-Pro-Administrator-Guide/Architecture
11. IBM QRadar SIEM Product Documentation, IBM. URL: https://www.ibm.com/docs/en/qsip/7.4?topic=deployment-qradar-architecture-overview

12. Darktrace Enterprise Immune System Technical Overview. URL: https://www.darktrace.com/news/darktrace-launches-enterprise-immune-system-version-4
13. FireEye Helix Platform Guide, FireEye Inc. URL: https://docs.fireeye.com/docs/landing.html
14. Fortinet FortiAI: Product Sheet, Fortinet. URL: https://www.fortinet.com/solutions/enterprise-midsize-business/fortiai
15. CrowdStrike Falcon Platform Overview. URL: https://www.crowdstrike.com/en-us/blog/driving-cybersecurity-forward-new-innovations-fal-con-2024/
16. AI in Cybersecurity – Research Report, MIT Technology Review. URL: https://www.technologyreview.com
17. GDPR – General Data Protection Regulation. URL: https://gdpr.eu/