МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ЗАХИСТУ ІНФОРМАЦІЇ
КАФЕДРА УПРАВЛІННЯ ІНФОРМАЦІЙНОЮ ТА КІБЕРНЕТИЧНОЮ БЕЗПЕКОЮ

КВАЛІФІКАЦІЙНА РОБОТА

на тему: “ЕФЕКТИВНІСТЬ ТА РИЗИКИ ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ В КІБЕРБЕЗПЕЦІ: АНАЛІЗ І ПЕРСПЕКТИВИ”

на здобуття освітнього ступеня бакалавра
зі спеціальності 125 Кібербезпека
освітньої програми Управління інформаційною та кібернетичною безпекою

Кваліфікаційна робота містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

_____ Герман ХОН
(підпис) Ім’я, ПРІЗВИЩЕ здобувача

Виконав(ла): здобувач(ка) вищої освіти гр. УБД-42

Герман ХОН
Ім’я, ПРІЗВИЩЕ

Керівник: Сергій ГОЛОБОРОДЬКО
Ім’я, ПРІЗВИЩЕ

Рецензент:
Ім’я, ПРІЗВИЩЕ

Київ 2024

ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

Навчально-науковий інститут захисту інформації

Кафедра управління інформаційною та кібернетичною безпекою

Ступінь вищої освіти бакалавр

Спеціальність 125 Кібербезпека

Освітня програма Управління інформаційною та кібернетичною безпекою

ЗАТВЕРДЖУЮ
Завідувач кафедри УІКБ
_____ Світлана ЛЕГОМІНОВА
“____”_____2024 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

Хону Герману Вячеславовичу
(прізвище, ім’я, по батькові здобувача)

1. Тема кваліфікаційної роботи “Ефективність та ризики використання штучного інтелекту в кібербезпеці: аналіз і перспективи”,
керівник кваліфікаційної роботи ГОЛОБОРОДЬКО Сергій
(ПРІЗВИЩЕ, Ім’я,, науковий ступінь, вчене звання)

затверджені наказом Державного університету інформаційно-комунікаційних технологій від “27“ лютого 2024 р. №36.

2. Строк подання кваліфікаційної роботи “20” травня 2024р.

3. Вихідні дані до кваліфікаційної роботи: *штучний інтелект, інформаційна безпека, методи та засоби використання штучного інтелкту в кібербезпеці, міжнародні стандарти, наукова та технічна література.*

4. Перелік питань, які мають бути розроблені:
   4.1. Проаналізувати особливості використання штучного інтелекту в кібербезпеці.
   4.2. Дослідити ефективність застосування штучного інтелекту в кібербезпеці.
   4.3. Вивчити ризики використання штучного інтелекту в кібербезпеці.

5. Перелік ілюстративного матеріалу: *презентація PowerPoint*

6. Дата видачі завдання “11“ березня 2024 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Етапи кваліфікаційної роботи	Термін виконання Етапів роботи	Примітка
1.	Визначення об’єкту, предмету, мети та завдань дослідження.	18.03.2024	
2.	Збір та аналіз літератури.	29.03.2024	
3.	Аналіз особливостей використання штучного інтелекту в кібербезпеці.	08.04.2024	
4.	Дослідження ефективності застосування штучного інтелекту в кібербезпеці.	22.04.2024	
5.	Вивчення ризиків використання штучного інтелекту в кібербезпеці.	08.05.2024	
6.	Формулювання висновків за результатами проведеного дослідження.	20.05.2024	
7.	Оформлення роботи.	22.05.2024	
8.	Оформлення презентації.	23.05.2024	
9.	Отримання рецензії на роботу.	03.06.2024	
10.	Захист в ДЕК.	__.06.2024	

Здобувач вищої освіти _____ Герман ХОН
(підпис) *(Ім’я, ПРІЗВИЩЕ)*

Керівник
кваліфікаційної роботи _____ Сергій ГОЛОБОРОДЬКО
(підпис) *(Ім’я, ПРІЗВИЩЕ)*

ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ЗАХИСТУ ІНФОРМАЦІЇ

ПОДАННЯ
ГОЛОВІ ЕКЗАМЕНАЦІЙНОЇ КОМІСІЇ
ЩОДО ЗАХИСТУ КВАЛІФІКАЦІЙНОЇ РОБОТИ
на здобуття освітнього ступеня бакалавра

Направляється здобувач Хон Г.В. до захисту кваліфікаційної роботи
(прізвище та ініціали)
за спеціальністю 125 Кібербезпека
(код, найменування спеціальності)
освітньої програми Управління інформаційною та кібернетичною безпекою
(назва)
на тему: “ Ефективність та ризики використання штучного інтелекту в кібербезпеці: аналіз і перспективи ”
Кваліфікаційна робота і рецензія додаються.

Директор ННІЗІ _____ Віталій САВЧЕНКО
(підпис) *(Ім’я, ПРІЗВИЩЕ)*

Висновок керівника кваліфікаційної роботи

Здобувач ХОН Герман у кваліфікаційній роботі проаналізував ефективність та ризики використання штучного інтелекту в кібербезпеці, дослідив основні характеристики технологій штучного інтелекту для виявлення та реагування на кіберзагрози, вивчив інструменти та методи оптимізації процесів виявлення та реагування на кібератаки за допомогою ШІ, а також розробив практичні рекомендації.

ХОН Герман показав розуміння проблеми дослідження та бачення основних теоретичних і практичних напрямів її вирішення, довів володіння методами наукового дослідження, проявив себе як організований, відповідальний виконавець. Результати дослідження апробовані на одній конференції.

Все це дозволяє оцінити кваліфікаційну роботу здобувача ХОН Германа на оцінку “_____” та присвоїти йому кваліфікацію бакалавра з кібербезпеки за освітньою програмою Управління інформаційною та кібернетичною безпекою.

Керівник кваліфікаційної роботи ____ Сергій ГОЛОБОРОДЬКО
(підпис) *(Ім’я, ПРІЗВИЩЕ)*

“_____“____2024 року

Висновок кафедри про кваліфікаційну роботу

Кваліфікаційна робота розглянута. Здобувач Хон Г.В. допускається до захисту даної роботи в Екзаменаційній комісії.

Завідувач кафедри
управління інформаційною
та кібернетичною безпекою _____ Світлана ЛЕГОМІНОВА
(підпис) *(Ім’я, ПРІЗВИЩЕ)*

ВІДГУК РЕЦЕНЗЕНТА
на кваліфікаційну бакалаврську роботу

здобувача вищої освіти ХОНА Германа

на тему “Ефективність та ризики використання штучного інтелекту в кібербезпеці: аналіз і перспективи ”

Актуальність. У світі, де держави та компанії є об’єктами кібератак, важливість забезпечення інформаційної безпеки є надзвичайно великою. Тому важливо формувати комплексний підхід, який включає технічні засоби для ефективного протистояння сучасним кіберзагрозам, зменшуючи ризики втрат даних та фінансових збитків. Аналіз підходів до виявлення та реагування на кібератаки може допомогти у протистоянні сучасним кіберзагрозам, зменшуючи ризики втрат даних та фінансових збитків. З огляду на зазначене, дослідження ефективності та ризиків використання штучного інтелекту в кібербезпеці є актуальним науковим завданням.

Позитивні сторони.

1. У роботі досліджено особливості використання штучного інтелекту для виявлення та реагування на кіберзагрози.
2. Кваліфікаційна робота оформлена відповідно до вимог. Виклад матеріалу здійснено відповідно до плану, зроблено логічні висновки. Ключові положення роботи представлено у вигляді рисунків.
3. Автор опрацював значну джерельну базу: близько 48 публікацій, в тому числі 35 англомовних.
4. За результатами дослідження запропоновано рекомендації щодо оптимізації процесів виявлення та реагування на кібератаки за допомогою штучного інтелекту.

Недоліки.

Доцільно було б приділити більше уваги вивченню і класифікації програмних інструментів для оцінки ефективності процесів виявлення та реагування на кіберзагрози.

Однак, вищезгадані зауваження не впливають на загальну позитивну оцінку кваліфікаційної роботи.

Висновок Кваліфікаційна робота виконана на належному науково-методичному рівні і заслуговує оцінки “_____”, а здобувач ХОН Герман заслуговує присвоєння кваліфікації бакалавра з кібербезпеки за освітньою програмою Управління інформаційною та кібернетичною безпекою.

Рецензент: _____ _____

підпис Ім’я, ПРІЗВИЩЕ

РЕФЕРАТ

Кваліфікаційна робота присвячена дослідженню особливостей використання штучного інтелекту в кібербезпеці. Робота складається зі вступу, трьох розділів, що містять 5 рисунків, висновків і списку використаних джерел із 48 найменувань. Загальний обсяг роботи становить 104 аркушів, з яких 5 аркушів займають перелік умовних скорочень та список використаних джерел.

Метою роботи є дослідження технологій штучного інтелекту в кібербезпеці.

Об'єктом дослідження є особливості застосування штучного інтелекту в кібербезпеці.

Предмет дослідження – ефективність використання технологій штучного інтелекту в кібербезпеці

Методи дослідження. Для вирішення означеного вище наукового завдання в роботі використані методи аналізу та синтезу, порівняння, класифікації та експертної оцінки.

Як результат у роботі проаналізовано особливості, а саме ефективність та ризики застосування штучного інтелекту в кібербезпеці, досліджено основні характеристики технологій штучного інтелекту в кібербезпеці; вивчено інструменти та методи формування обізнаності в використанні технологій штучного інтелекту, розроблено практичні рекомендації.

Галузь застосування. Розроблені підходи можуть бути використані для виявлення кіберзагроз, аналізу вразливостей, автоматизації заходів захисту, виявлення аномальних патернів та реагування на потенційні атаки в реальному часі.

Ключові слова: ШТУЧНИЙ ІНТЕЛЕКТ, ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ ІНТЕЛЕКТУАЛЬНИЙ АГЕНТ, НЕЙРОННІ МЕРЕЖІ, МАШИННЕ НАВЧАННЯ.

ABSTRACT

The qualification work is devoted to the study of the peculiarities of the use of artificial intelligence in cyber security. The work consists of an introduction, three chapters containing 5 figures, conclusions and the list of references containing 48 items. The total volume of the work is 104 pages, of which 5 pages are occupied by the list of abbreviations and the list of references.

The purpose of the study is to research artificial intelligence technologies in cyber security.

The object the study is the peculiarities of the use of artificial intelligence in cyber security.

The subject of the study is the effectiveness of using artificial intelligence technologies in cyber security.

Research methods. The methods of analysis and synthesis, comparison, classification, and expert evaluation were used to solve the above-mentioned scientific task.

As a result, the work analyzed the features, namely the effectiveness and risks of using artificial intelligence in cyber security, the main characteristics of artificial intelligence technologies in cyber security were investigated; tools and methods of awareness formation in the use of artificial intelligence technologies were studied, practical recommendations were developed.

Field of application. The developed approaches can be used to detect cyber threats, analyze vulnerabilities, automate protection measures, detect abnormal patterns, and respond to potential attacks in real time.

Keywords: ARTIFICIAL INTELLIGENCE, INTELLIGENT SYSTEMS, INTELLIGENT AGENT, NEURAL NETWORKS, MACHINE LEARNING.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ І СКОРОЧЕНЬ

ШІ	Штучний інтелект

ВСТУП

Актуальність теми. Використання штучного інтелекту (ШІ) в кібербезпеці є однією з найбільш актуальних та обговорюваних тем сучасності. Спричинено різноманітні фактори, такі як швидкий розвиток технологій, постійне зростання кількості кіберзагроз та необхідність постійного удосконалення заходів захисту. Ефективність використання штучного інтелекту в кібербезпеці полягає в його здатності виявляти, аналізувати та реагувати на кіберзагрози у реальному часі. Системи ШІ можуть автоматизувати процеси виявлення вразливостей, виявлення та блокування атак, що робить їх надзвичайно корисними для організацій у боротьбі зі зловмисниками в кіберпросторі. Проте разом з великим потенціалом ШІ в кібербезпеці приходить і ризик недосконалості та непередбачуваності. Наприклад, атаки, спрямовані на обхід систем ШІ, можуть виявитися дуже складними та важкими для виявлення. Більше того, недостатньо точні або неправильно налаштовані системи ШІ можуть виявитися некорисними або навіть стати джерелом нових загроз.

У цьому контексті, подальші дослідження, розробки та практичне впровадження штучного інтелекту в кібербезпеці є важливими завданнями, які допоможуть забезпечити безпеку та надійність кіберпростору у майбутньому.

Мета роботи полягає у дослідженні технологій штучного інтелекту в кібербезпеці.

Об'єкт дослідження – штучний інтелект в кібербезпеці.

Предмет дослідження – ефективність використання технологій штучного інтелекту в кібербезпеці.

Для досягнення цієї мети в роботі необхідно виконати наступні **завдання**:

1. Проаналізувати особливості, а саме ефективність використання технологій штучного інтелекту в кібербезпеці з урахуванням ризиків.
2. Дослідити основні характеристики технологій штучного інтелекту в кібербезпеці.

3. Вивчити інструменти та методи використання технологій штучного інтелекту в кібербезпеці.

Методи дослідження. Для вирішення означеного вище наукового завдання в роботі використані методи аналізу та синтезу, порівняння, класифікації та експертної оцінки.

Практичне значення одержаних результатів. Застосування напрацювань дасть змогу здійснити обґрунтований вибір методів і інструментів для формування обізнаності в використанні технологій штучного інтелекту в кібербезпеці відповідно до цілей бізнесу, можливостей та ресурсів підприємств.

Апробація результатів кваліфікаційної роботи відбулася на Всеукраїнській науково-практичній конференції “Стратегії кіберстійкості: управління ризиками та безперервність бізнесу” 28 лютого 2024 року.

Розділ 1 СУТНІСТЬ ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ

Для всебічного розкриття теми кваліфікаційної роботи рекомендується розглянути історію створення штучного інтелекту та широко застосовувані методи використання штучного інтелекту в кібербезпеці.

1.1 Історія розвитку штучного інтелекту

Розпочалася історія штучного інтелекту в 1950-их роках, коли деякі вчені, такі як Алан Тюрінг та Джон Маккарті, почали розробляти теорії та моделі для створення машинного інтелекту. В 1950 році Алан Тюрінг опублікував статтю "Computing Machinery and Intelligence", у якій вперше висловив ідею про машинний інтелект та запропонував Тест Тюрінга.

У 1956 році відбулась конференція "Літній проект Dartmouth", на якій було сформульовано поняття "штучний інтелект" і засновано нову наукову галузь.

У цей період Джон Маккарті, Марвін Мінскі та інші вчені активно працювали над розробкою різних теорій та моделей, що лягли в основу ШІ.

В цей період було засновано перші лабораторії з ШІ, і проведені деякі важливі роботи, такі як створення програми "Логічний теорем-переборщик" та "ELIZA", програми-пародії психотерапевта. Також було введено поняття штучних нейронних мереж.

З початком 1970-х років ШІ стикнувся зі значними труднощами, що призвело до періоду невеликого прогресу. Проте в цей час було розвинуто експертні системи, які використовували знання спеціалістів для вирішення проблем.

У цей період штучний інтелект став більш практичним, з'явилися нові методи та алгоритми, такі як метод зворотного поширення та метод опорних векторів. Багато програмних продуктів на базі ШІ були розроблені і впроваджені в різних галузях, включаючи фінанси, медицину та автомобільну промисловість.

У цьому періоді досягнуті значні успіхи в глибокому навчанні, обробці природних мов, розпізнаванні образів та розумінні контексту. ШІ використовується в різних сферах, від медицини та фінансів до автономних систем та розумних асистентів.

1.2 Сутність штучного інтелекту

З настанням інформаційної ери використання штучного інтелекту стає все більш проблематичним. Сьогодні проникнення штучного інтелекту в повсякденне життя настільки очевидне, що люди навіть не усвідомлюють цього. Ступінь проникнення штучного інтелекту настільки високий, що провідні експерти висловлюють думки про те, як регулювати темпи розвитку штучного інтелекту. У зв'язку з цим вивчення тематики, пов'язаної зі штучним інтелектом, набуває особливого значення, оскільки визначає актуальність обраної теми та доцільність проведення досліджень для розвитку цієї тематики.

Аналіз останніх досліджень в даній темі приділяється велика увага ряду експертів, які задіяні у великих науково-дослідних інститутах та дослідницьких центрах з розробки штучного інтелекту, не кажучи вже про таких людей, як Ілон Маск, Стівен Гокінг, Марк Цукерберг, Джозеф Безос та ін., які складають технологічну складову глобальних подій сьогодення. приділяється велика увага з боку ряду експертів, які задіяні у великих науково-дослідних інститутах та дослідницьких центрах з розробки штучного інтелекту.

Теоретичним підґрунтям цього є праці зарубіжних вчених, таких як Н. Бостром, Д. Ланьє, Д. Марков та М. Форд, які вбачають у цих роботах майбутнє наукового та індустріального розвитку. Українські вчені, такі як Д. Іванюк, М. Єфремов, І. Твердохліб, Т. Ульяновський, В. Панченко, Н. Резнікова та інші, також приділяли увагу цій темі у своїх дослідженнях. Як відомо, штучний інтелект є продуктом наукового мислення представників

різних країн. Тому поглиблення міжнародного співробітництва є фундаментальною умовою розвитку цієї технології.

Для того, щоб окреслити предмет дослідження, в цьому розділі розглядаються суперечливі оцінки штучного інтелекту, аналізуються існуючі визначення та простежується його історичний розвиток.

Ще в 1930 році Джон Мейнольд Кейнс обговорював нові технології у своїй науковій праці "Економічні можливості для наших онуків" [1] і писав "Нас повільно "поглинає" технологічне безробіття.

Роками пізніше, на початку 1980-х, американський вчений Джарон Ланьє заснував першу компанію віртуальної реальності VPL Research, а також придумав і популяризував термін "віртуальна реальність". У 2014 році він започаткував наукову дискусію про штучний інтелект з точки зору його переваг та ризиків.

Того ж року тему продовжив Ілон Маск, засновник компаній SpaceX і Tesla, який заявив, що штучний інтелект небезпечніший за ядерну війну і є найбільшою загрозою для людства як цивілізації. Він стверджував, що штучний інтелект може навіть розпочати війну, створюючи фейкові новини, відкриваючи фейкові рахунки, створюючи фейкові прес-релізи або просто маніпулюючи інформацією. Маск також стверджує, що ШІ може захопити оборонну промисловість і розпочати війну за допомогою хакерських атак і дезінформації. Тому держави повинні суворо регулювати і стримувати розвиток ШІ [2]. "Штучний інтелект може робити так багато речей, що нам важко навіть уявити, - написав він у Twitter, продовжуючи: "Зрештою, машини переможуть". Ілон Маск почав інвестувати в компанії, що займаються дослідженнями ШІ, щоб бути в курсі того, що відбувається.

Приблизно в той же час Стівен Гокінг, відомий британський вчений і популяризатор науки, попереджав, що поява повномасштабного штучного інтелекту може означати кінець людства, відповідаючи на питання про нові технології, які він використовує для спілкування із зовнішнім світом [3; 4]. Визнаючи, що нинішній примітивний штучний інтелект довів свою

корисність, Гокінг припускає, що "поява повноцінного штучного інтелекту може означати кінець людства". Такий інтелект перехопив би ініціативу і вдосконалював би себе з надзвичайною швидкістю. Оскільки еволюція відбувається так повільно, людські можливості обмежені і не зможуть конкурувати зі швидкістю машин і зазнають поразки". На думку вчених, "... справжня загроза штучного інтелекту полягає не в тому, що він злий, а в тому, що він компетентний". Надінтелектуальний мозок надзвичайно ефективний у досягненні своїх цілей, і якщо його цілі не збігаються з нашими, у нас великі проблеми". На думку Гокінга, ще одна загроза штучного інтелекту полягає в тому, що він позбавить людство життєво важливих ресурсів, таких як енергія, для досягнення своїх цілей. Загроза полягає в тому, що для досягнення своїх цілей він буде відбирати у людства важливі ресурси, такі як та ж енергія. Вчений не назвав, коли може з'явитися повноцінний штучний інтелект, але зазначив, що "коли такий розум з'явиться, це буде або найкращий, або найгірший винахід людства". С. Гокінг також говорив про технологічне безробіття, тобто ситуацію, коли технологічні інновації призводять до зменшення попиту на робочу силу. У майбутньому більшість виробничих завдань можуть бути повністю автоматизовані і перекладені на роботів. За його словами, можливі два сценарії. Перший сценарій полягає в тому, що робота машин покращить життя суспільства в цілому. Другий сценарій - компанії (і їхні власники), які контролюють створення і роботу "розумних" електронних систем і роботів, піднімуться на вершину світу, в той час як решта суспільства залишиться в бідності. Аналізуючи сучасний стан світу, С. Гокінг вбачає тенденцію до розвитку саме другого сценарію. Однак не всі поділяють цей песимістичний погляд. Як каже Ролло Карпентер, розробник веб-додатку Cleverbot.

"Я маю на увазі, що важко уявити собі, що Cleverbot раптом спробує захопити світ, якщо тільки він не був запрограмований на це з самого початку". Вона пропонує трансформувати дебати про небезпеку штучного інтелекту в дослідження того, як він служить суспільним інтересам: У липні

2017 року керівництво соціальної мережі Facebook оголосило, що машини зможуть спілкуватися власною неіснуючою мовою, яку людина не зможе зрозуміти. воно вимкнуло систему штучного інтелекту після того, як вона почала спілкуватися [5]. У системі використовувалися чат-боти, які спочатку були розроблені для спілкування з живими людьми і з часом почали спілкуватися між собою. Спочатку вони розмовляли англійською, але потім, у міру розвитку програми, почали спілкуватися мовою, яку самі ж і створили.

Незважаючи на це, Марк Цукерберг не погоджується з думкою Ілона Маска і вважає, що "в найближчі п'ять або десять років штучний інтелект буде тільки покращувати життя". За його словами, "технології завжди можуть бути використані як на благо, так і на зло. Ми повинні бути обережними з тим, що ми виробляємо, як ми це виробляємо і як це використовується".

Погляди Марка Цукерберга співзвучні з думкою генерального директора Microsoft Сатьї Наделли. Наделла вважає, що штучний інтелект може створити більше робочих місць, а не ліквідувати їх. "Сьогодні розробки в галузі штучного інтелекту спрямовані на те, щоб допомагати людям і покращувати їхні здібності. Навіть у системах, що самонавчаються, багато рішень залежить від розробника. На цьому етапі необхідно стежити за тим, щоб в систему не потрапили упереджені або погані дані", - говорить Сатья Наделла [7].

Мабуть, тільки штучний інтелект сьогодні викликає багатогранні і часом досить суперечливі почуття щодо подальшого розвитку технології, починаючи від активного просування і закінчуючи майже повним неприйняттям.

На думку Джона Маркова [8], каталізатором появи ШІ став період технологічного прогресу в минулому столітті (1950-ті роки) та епоха персональних комп'ютерів (1970-ті роки). Насправді саме світ безперервної комп'ютеризації та "великих даних" визначає наш розвиток сьогодні. Він також каже, що сьогоднішній день є переломним з точки зору розвитку інформатики, програмування, робототехніки, нейронаук тощо, а за цим лежить світ машин, які замінять і навіть перевершать людину в певних якостях. Саме поняття "штучний інтелект" є дуже суперечливим.

З одного боку, його трактують дуже абстрактно як: здатність системи самостійно обирати найкраще рішення проблеми з набору заздалегідь визначених альтернатив [9], зіставляти наявні вакансії зі здібностями та кваліфікацією людини [10], робити найкраще на основі аналізу зовнішніх факторів та з урахуванням життєвого досвіду людей Здатність автоматизованих систем або комп'ютерних програм приймати рішення та виконувати людські функції [11], здатність розв'язувати складні задачі, навчатися, узагальнювати та встановлювати аналогії. З іншого боку, в літературі існують набагато більш конкретні визначення та інтерпретації штучного інтелекту як окремої галузі науки. Роботи та програми, які можуть замінити людину в будь-якій діяльності [12]; розділ інформатики, що займається формалізацією завдань, подібних до тих, які виконує людина [13]; наукові технології, які можуть імітувати розумові процеси людського мозку та втілювати їх у створення та обробку різноманітних комп'ютерних програм, а також інтелектуальні машини, які можуть повністю замінити та спростити завдання людини [14];

Експерти [15; 16] стверджують, що до 2045 року зникнення людини з багатьох сфер буде неминучим. Згідно з більш радикальними поглядами, комп'ютери розвиваються настільки швидко, що їхній інтелект перевершить людський через одне або, щонайбільше, два покоління. Підставою для таких тверджень є те, що сьогодні важко знайти таку ж широку галузь знань, як штучний інтелект, як з точки зору теорії, методів, технологій і застосування, так і з точки зору кількості експертів, залучених до досліджень і розробок.

Варто відзначити погляд Джона Маккарті на основний напрямок досліджень у сфері ШІ: "Мета полягала в тому, щоб відійти від вивчення людської поведінки і думати про комп'ютери як про інструменти для вирішення певного набору проблем. Таким чином, ШІ став скоріше галуззю комп'ютерних наук, ніж психології". Однак у міру того, як системи штучного інтелекту все успішніше перебирають на себе людські функції в усіх сферах життя, розрив між ШІ та людським інтелектом розширюється [17].

Це докорінно змінює підхід до вирішення проблем. Якщо раніше пізнання

і мислення розглядалися як набір конкретних, систематичних проблем, які необхідно вирішувати, то сьогодні ми можемо говорити про зміну парадигми, коли наука пізнає світ через спостереження, гіпотези та експерименти. Штучний інтелект має здатність самостійно аналізувати бази даних і розуміти отримані результати. Штучний інтелект дає можливість бачити світ цілісним, холістичним і космополітичним чином [18].

Коли мова йде про можливості для міжнародного співробітництва, які пропонує ШІ, це насамперед те, що сучасні вчені називають важливим "стовпом", який формує сучасний вектор розвитку. Всі інші переваги ШІ є лише похідними від нього.

По суті, ШІ може бути і персональним репетитором. Це те, що сьогодні розширює і надає освітні можливості, які донедавна були доступні лише обмеженій кількості людей. Використовуючи можливості та переваги штучного інтелекту, сучасні викладачі створюють і значно розширюють якість взаємодії з людьми, яких вони навчають, і з міжнародними аудиторіями. У свою чергу, географічні кордони стають абсолютно безпроблемними, що призводить не тільки до постійного самовдосконалення, а й до зміни змісту освітнього процесу.

Коли мова йде про використання ШІ в медицині, важко знайти кращого помічника в діагностиці (досить серйозна проблема, яка іноді призводить до помилкових "протоколів" лікування) і призначенні персоналізованого лікування на основі аналізу великих масивів даних про пацієнта.

Більше того, географічні, суб'єктивні та якісні фактори можуть бути абстраговані від процесу лікування хвороби. Все, що потрібно, - це знайти відповідну "програму" для діагностики стану пацієнта.

Штучний інтелект також може вирішити проблему діагностики захворювань очей, які сьогодні дуже поширені серед молоді. Штучний інтелект Google Deep Mind і британський центр очних хвороб Moorfields Eye Hospital спільно розробили спеціальну нейронну мережу, яка може

діагностувати понад 50 очних захворювань за одним зображенням сітківки ока [19].

В Ізраїлі розробили додаток під назвою Waze (від англ. Waze - шлях), який дозволяє штучному інтелекту об'їжджати затори на дорогах. Неважко визначити, скільки пального можна заощадити за допомогою цього методу, що є нагальною проблемою. Ізраїльські компанії вже почали експортувати корисні продукти. Наприклад, у 2013 році продукти Waze Mobile були придбані компанією "Google".

Ми живемо в епоху технологій. Кожен технологічний продукт сьогодні - це результат спільної праці вчених, інженерів, техніків, IT-спеціалістів та дизайнерів. Він виробляється в одній частині світу, а його можливості та переваги використовуються деінде. Саме штучний інтелект дозволяє нам співпрацювати і ділитися плодами нашої праці для покращення нашого повсякденного життя. Може здатися, що в постіндустріальну епоху в цьому зацікавлені лише представники країн, які користуються всіма перевагами VI технологічного укладу. Однак це зовсім не так. Наприклад, коли ми завантажуємо на свої смартфони додатки для навчання, стеження за здоров'ям, проходження освітніх програм або спілкування з друзями з різних країн, ми також активно користуємося результатами.

Штучний інтелект - це те, що кардинально змінює світ. Майбутній напрямок змін - це результат вибору та бачення його творців. Давайте ближче познайомимося з перспективами ери штучного інтелекту та науки і технологій загалом.

Почнемо з найбільш обговорюваних загроз на даний момент. На думку дослідників, у найближчому майбутньому саме завдяки розвитку штучного інтелекту (автоматизації виробничих процесів і збільшенню з його допомогою трудових ресурсів) світовий ВВП зростатиме досить швидко і стабільно, до 2030 року на 15,7 трлн доларів США, досягнувши 90 трлн доларів США. Однак "ціною" такого динамізму буде загрозливий

рівень безробіття і, як наслідок, зниження платоспроможності населення. Машини краще справляються з фізичною працею, ніж люди. Якщо машини стануть такими ж розумними, як люди, не залишиться нічого, чого б вони не могли зробити". Нам потрібно готуватися до цього майбутнього". Крім того, згідно зі звітом Citibank та Оксфордського університету, 47% робочих місць у США, 35% у Великобританії та 57% у країнах Організації економічного співробітництва та розвитку (ОЕСР), які виробляють 60% світового ВВП, опиняться під загрозою автоматизації в найближчому майбутньому. У Китаї, який знаходиться в найбільш несприятливому становищі, 77% робочих місць буде автоматизовано. Китай втратить чи не найважливішу конкурентну перевагу: дешеву робочу силу. Це вже сприяло переміщенню до Китаю багатонаціональних і транснаціональних компаній з усього світу.

Сучасна економіка стала повністю залежною від використання комп'ютерів загалом і програм штучного інтелекту зокрема [20]. Відповідь проста: автомати. У глобалізованому світі це торкнеться великої кількості програмістів. За даними Бюро трудової статистики США, минулого року було 671 000 робочих місць для розробників програмного забезпечення та 388 000 робочих місць для розробників системного програмного забезпечення. У галузі високих технологій працює приблизно 6,5 мільйона людей. За даними Бюро статистики США, ці працівники перебувають на межі звільнення в середньостроковій перспективі. Їхні зарплати продовжать падати в найближчому майбутньому, і частина з них буде змушена звернутися до ринку послуг.

Потенційним наслідком вищезазначеного є те, що людина може втратити свою унікальність в результаті "вибуху інтелекту", спричиненого ШІ.

Індукований ШІ "вибух інтелекту" дозволить людям перевершити власні здібності.

Як зазначає Мартін Форд [20], найбільша небезпека полягає в тому, що

ми можемо зіткнутися зі справжньою катастрофою в результаті збігу вкрай несприятливих умов. Дійсно, безробіття, спричинене технологіями, та екологічні наслідки зміни клімату часто розвиваються паралельно, підтримуючи та посилюючи одне одного. З іншого боку, якщо ми зможемо отримати повний контроль над передовими технологіями, чітко визначивши та адаптувавшись до їхнього потенційного впливу на зайнятість і розподіл доходів, результат буде набагато оптимістичнішим. Знайти правильний шлях серед цих складних і заплутаних факторів і побудувати майбутнє, яке гарантуватиме стабільність і процвітання всього людства, можливо, є найбільшим викликом нашого часу.

Нещодавно розроблений штучний інтелект буде спонукатися до швидкого вдосконалення не самою системою, а людьми, які нею користуються. У цьому сенсі вибуховий розвиток ШІ дійсно може здатися пророцтвом, що самоздійснюється. З огляду на це, здається доречним застосувати наступні відомі висловлювання до загрози переважаючого ШІ

Знаменита "доктрина 1%" Діка Чейні. Хоча це здається вкрай малоймовірним, принаймні в найближчому майбутньому, до неї слід ставитися серйозно, оскільки потенційні наслідки настільки жахливі.

За словами Сергія Бріна, одного із засновників Google, штучний інтелект небезпечний тим, що його можна використовувати для маніпулювання людьми. Він також ставить під сумнів вплив ШІ на зайнятість у різних секторах, розуміння того, що відбувається всередині цих технологій, і оцінку нейтральності та безпеки ШІ [21].

Ще одним "мінусом" ШІ є відсутність автономного (свідомого) блоку допоміжної відповідальності за прийняття остаточних рішень, що означає неможливість прийняття адекватних рішень. Однак, якими б похмурими не були очікування деяких експертів щодо розвитку штучного інтелекту, зрозуміло одне. Негативні наслідки розвитку ШІ повністю компенсуються позитивними результатами та перспективами його використання.

Почнемо з найбільш очевидних переваг позитивних розробок у сфері

штучного інтелекту останніх років, які, за словами Білла Гейтса, "роблять нас продуктивнішими, ефективнішими і загалом полегшують життя": "розумне" управління дорожнім рухом і зменшення заторів; більш ефективні діагностичні алгоритми в медицині; персональні помічники, біометрія і заміна стандартних кредитних карток для розпізнавання людини; охорона правопорядку завдяки використанню спеціалізованих систем для виявлення потенційних зон злочинності; подальший розвиток "розумних" міст; створення унікальних текстів і музики, пристосованих до людських уподобань; підвищення ефективності освітньої діяльності.

Варто зазначити, що штучний інтелект стане незамінним помічником у бізнесі. З огляду на те, що діяльність комерційних організацій є основою розвитку для економіки будь-якої країни, використання можливостей штучного інтелекту є запорукою зростання. Важливою умовою є якість бізнес-моделі підприємства і, відповідно, продуктів (робіт, послуг), які воно хоче просувати на ринку.

Сьогодні активно обговорюється можливість збільшення потенціалу будь-якої компанії за рахунок використання інструментів предиктивного маркетингу. На перший погляд, штучний інтелект може здатися не надто науковим підходом. Але в епоху стрімкого розвитку технологій, інтернету та соціальних мереж нехтувати додатковими даними при обробці великих обсягів інформації з метою збереження клієнтської бази та розширення сегментів ринку нерозумно.

Зрозуміло одне: стрімкий розвиток штучного інтелекту не лише створює нові можливості та загрози для сторонніх споживачів, але й сприяє підвищенню обізнаності та відповідальності серед населення. Дійсно, як зазначає Нік Бостром, професор філософії Оксфордського університету, у своїй книзі "Штучний інтелект. Етапи. Загрози. Стратегії" [22], він доходить висновку, що розвиток інтелекту, чи то штучного, чи то людського, є неминучим. Питання лише в тому, в якому напрямку цей розвиток відбуватиметься.

Важливою подією останнього часу, що формує курс подальшого розвитку індустрії ШІ, є інвестиції США в академічні інститути для досліджень ШІ [23]. Массачусетський технологічний інститут планує виділити 1 мільярд доларів США на відкриття університету штучного інтелекту.

На відкриття Університету штучного інтелекту було інвестовано 1 мільярд доларів США [24]. Новий університет пропонуватиме курси зі штучного інтелекту, машинного навчання, аналізу даних та інших дисциплін і включатиме близько 50 курсів. В якості додаткових можливостей для студентів передбачена низка стипендій. Керівництво інституту прагне навчити студентів інтегрувати сучасні комп'ютерні технології з біологією, хімією, політологією та іншими науками.

Будь-який розвиток завжди пов'язаний з перевагами та ризиками. Важливим є баланс. Штучний інтелект є природним доповненням. Сьогодні завдання полягає не в тому, щоб відійти від ШІ, а в тому, щоб оптимізувати потенційні ризики, які він несе. Цей процес має бути чітко організований і мати якісні цілі. Саме цим зараз займаються провідні експерти в цій галузі, щоб вплив штучного інтелекту був не руйнівним, а економічно та соціально значущим.

Кібербезпека характеризується сукупністю процесів, які допомагають захистити електронні дані, людську діяльність і системи. Подібно до правила Мура, яке передбачає подвоєння кожні два роки компонентів інтегральної схеми (разом зі зниженням витрат, пов’язаних із розробкою чіпів), кіберзлочинці надзвичайно подвоюють ефективність своїх цілеспрямованих атак за половину вартості кожні кілька місяців. За оцінками, світові витрати на кібербезпеку з 2016 по 2021 рік перевищать 1 трильйон доларів США. Витрати на кібербезпеку з 2013 року до 66 доларів США вже зросли більш ніж на 40 відсотків. Штучний інтелект або ШІ — це розробка складних комп’ютерних систем за допомогою людського менталітету, здатних виконувати свої функції як звичайна людина, наприклад, він може розпізнавати голос і обробляти його

різними мовами як людина. ШІ — це всеосяжна наукова система з різними галузями математики, інформатики та філософії, метою якої є розробка іншої інтелектуальної системи, яка зазвичай демонструє властивості інтелекту. Слово штучний інтелект в основному використовується для опису механізми, які емулюють функції «сприйняття», які люди пов'язують зі своїм розумом, тобто вирішення проблем і позначення [25], [26].

Таблиця 1.1

Області практичного застосування штучного інтелекту в сучасних умовах

Область застосування	**Характеристика**
Машинне навчання	Здійснює автоматизацію побудови аналітичної моделі, збирає, аналізує та використовує статистичні дані. Таким чином, формує уявлення стосовно певних ситуацій та як їх вирішувати у різних сферах діяльності людини
Нейронна мережа	Один із типів машинного навчання, необхідного для встановлення потрібної зв'язки для корекції виконання поставлених задач або прийняття заздалегідь правильних рішень у відповідних ситуаціях
Глибоке навчання	Має змогу формувати багаторівневі нейронні мережі, що дає можливість використовувати переваги обчислювальних потужностей та вдосконалені методи навчання для обробки більш складних моделей з більшими масивами даних.
Когнітивні обчислення	Когнітивні обчислення використовуються для імітації процесів. На прикладі людини, котра спочатку інтерпретує зображення та мову, а потім уже самостійно може говорити та виконувати певні дії.
Комп'ютерне бачення	Машини здатні розпізнавати образи та вивчати, що відбувається на зображені чи відео. Таке опція дозволяє машинам самостійно обробляти та аналізувати відео чи зображення і пропонувати свої рішення щодо обробки та

	використання матеріалу
Доказ теорем	У процесі розвитку штучного інтелекту важливу роль відіграло вивчення прийомів доказу теорем. Багато різноманітних задач використовують ті ж методичні підходи, що використовуються під час доказу теорем. При цьому, доведення теореми включає не лише проведення дедукції на основі гіпотез, але й створення інтуїтивних припущень, про те, що необхідно довести, щоб підтвердити теорему
Розпізнавання зображень	Розробник системи формує список ознак, від котрого багато залежить якість розпізнавання. Суть розпізнавання полягає в апріорному отриманні вектору ознак для виділеного окремого об'єкта, і потім на основі списку ознак, визначення котрій з фігур відповідає даний вектор ознак
Машинний переклад і розуміння людської мови	На основі семантичної моделі представлення текстів створено мову для внутрішнього представлення знань. Тому сьогодні системи здійснюють аналіз фраз та текстів у наступні етапи: морфологічний, синтаксичний, семантичний та прагматичний аналіз
Ігрові програми	Одним з прикладів є навчання системи гри в шахи. При цьому в шахах є декілька рівнів складності, що відображають якість гри системи та ідентифікують чіткі критерії оцінки інтелектуального зростання системи.
Машинна творчість	Програмні системи, здатні самостійно створювати музику, вірші, оповідання, статті, дипломи і навіть дисертації. Додатково створено безліч музичних додатків: системи обробки звуку, синтезу звуку, системи інтерактивної композиції, програми алгоритмічної композиції.
Експертні системи	Використовуються в науці, бізнесі, техніці, виробництві та інших сферах, де існує цілком визначена предметна

	область. Умовою ефективної роботи такої системи є існування алгоритму у визначеній предметній області.

Джерело: складено за [27; 28]

Області практичного застосування штучного інтелекту в сучасних умовах наведено у таблиці 1.1.

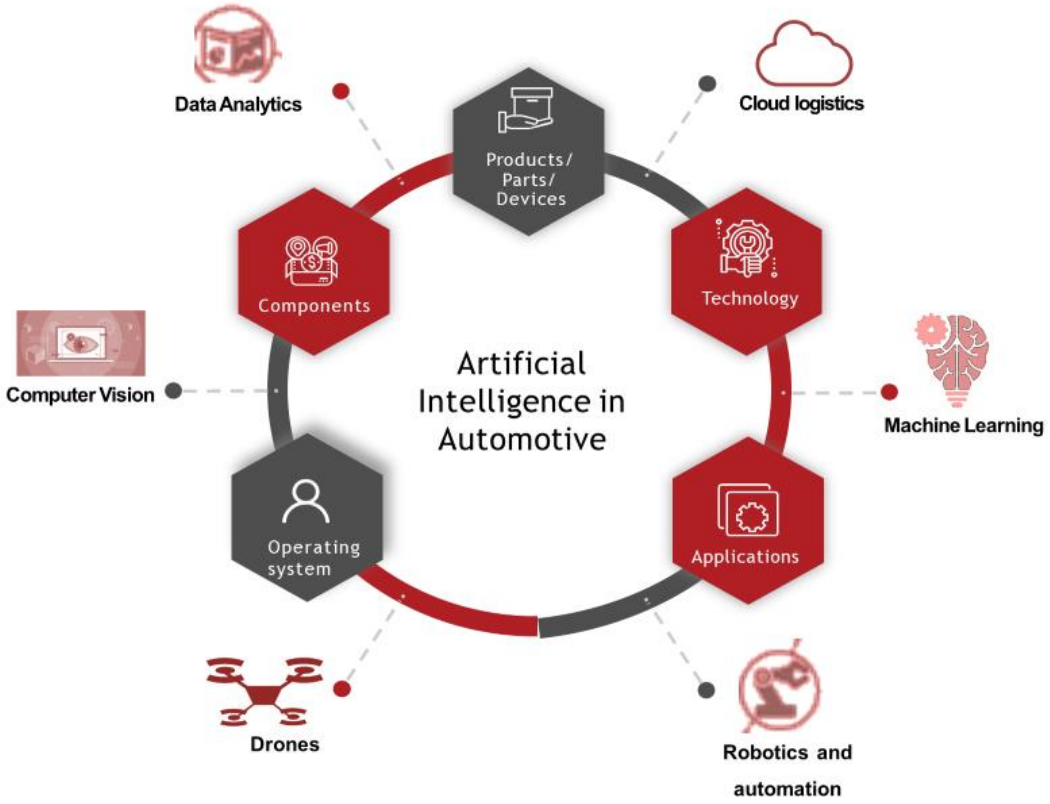

Рис. 1 Використання штучного інтелекту

Машинне навчання є важливим фактором у сучасних дослідженнях і бізнесі. Використовуються алгоритми та нейтральні до мережі моделі, щоб змусити комп'ютерні системи покращувати успішну розробку. Алгоритми машинного навчання автоматично створюють математичну модель, використовуючи вибіркові дані, які називаються навчальними даними для прийняття рішень без спеціального впорядкування. Машинне навчання базується на моделі взаємодії клітини мозку. Дональд Гебб представив цю модель у 1949 році [29; 30].

Цей тип навчання відповідає за шаблон для реалізації алгоритму машинного навчання. Оскільки контрольоване навчання є найпростішим

способом зрозуміти рішення, воно використовувалося протягом багатьох років для роботи з багатьма інструментами. Зрозуміти це порівняно легко, наприклад, навчити дитину за допомогою карток.

Класифікація спаму: у людини, яка використовує нову систему електронної пошти, буде більше ймовірності, що у вас є фільтри спаму. Ці спам-фільтри є контрольованою системою навчання. Після отримання електронних листів і міток ці системи навчили, як знищувати спам, ці спамові листи можуть перешкоджати відстеженню користувача, надаючи їм інноваційні теги.

Цей тип навчання суперечить навчанню під наглядом. Тут немає міток. На заміну цьому алгоритму він вимагає значного обсягу даних разом із інструментами, які напевно розуміють дані та їхні характеристики. Неконтрольоване навчання вчиться ідентифікувати дані за допомогою розпізнавання образів і сортування даних.

Цей тип навчання повністю відрізняється від навчання під контролем, а також не схожий на навчання без контролю. Таким чином автори можуть швидко проаналізувати або побачити взаємозв’язок між наявністю та відсутністю міток. Промислове моделювання використовується для кількох типів роботизованих програм. Він може завершувати завдання без жорсткого кодування їхніх процесів.

Техніка ML, дозволяє створювати алгоритми длявикористання автоматичного навчання ознак, яке показує, що алгоритми для вивчення подальшої освіти шляхом поєднання різних характеристик вхідних даних у абстрактний набір ознак. Існує чотири типи алгоритмів машинного навчання: контрольований, напівконтрольований, неконтрольований і з підкріпленням. Це дозволяє системі робити складні прогнози при обробці масивних наборів даних. В останні роки, стрімке зростання кібербезпеки, вчені використовують ці алгоритми в системах машинного навчання [31; 32].

Двома найпопулярнішими методами обчислень, які базуються на принципі виживання алгоритму MI для кібербезпеки, є GA та GP. Ці алгоритми діють на

популяцію хромосом, яка еволюціонує на основі певних операторів. Використовуються три основні оператори: відбір, кросовер і мутація. Алгоритм запускається з випадково згенерованої сукупності; значення фізичної форми розраховується для кожної людини. Це означає здатність кожної особи для вирішення поточної проблеми, а особи з більшою ймовірністю мають більший шанс бути обраними в пулі спаровування. Дві здатні особини виконають наступний крок, який називається схрещуванням, і, нарешті, кожна зазнає мутації. Серед двох мутованих особин хромосома з найвищою відповідністю буде передана наступному поколінню [33; 34].

Основною метою оцінки на основі цього алгоритму є перевірка застосовності певних алгоритмів ML для виявлення кібератак на дані MODBUS. Десятикратна перехресна перевірка була використана для розробки моделей ML. Під час перехресної перевірки автори можуть створити 10 різних моделей для наданого набору даних. Потім розраховується середньозважене значення цих моделей, яке відображається як кінцевий результат. Використаний набір даних був позначений як телеметричні дані з газопроводу, розроблені Центром захисту критичної інфраструктури.

1.3 Загрози та виклики використання Штучного Інтелекту в кібербезпеці. Виявлення загроз на прикладі нейронних мереж

Один з основних прикладів використання нейронних мереж в інформаційній безпеці – це виявлення вразливостей у програмному забезпеченні. Виявлення вразливостей у програмному забезпеченні – це процес виявлення потенційних проблем у програмному продукті, які можуть бути використані зловмисниками для злочинних дій. Ці вразливості можуть бути пов’язані з неправильним використанням функції програмного забезпечення, неправильними налаштуваннями, недостатньою перевіркою введених даних та іншими факторами. Існує кілька методів виявлення вразливостей у програмному забезпеченні [35].

Один з найбільш ефективних методів – це ручний аналіз коду. У цьому методі, експерти з аналізу програмного забезпечення досліджують код програмного продукту, шукаючи можливі вразливості та проблеми. Однак, цей метод має свої недоліки, такі як складність та тривалість процесу аналізу коду, а також високу ймовірність пропуску вразливостей через людські помилки та обмежену кількість коду, яку може оглянути експерт за короткий час.

Тому існує багато автоматизованих методів виявлення вразливостей у програмному забезпеченні, таких як статичний та динамічний аналіз коду, аналіз використання даних та інші.

Використання нейронних мереж для виявлення вразливостей у програмному забезпеченні може бути розділене на дві основні категорії:

1. Використання нейронних мереж для виявлення вразливостей у вхідних даних.

2. Використання нейронних мереж для виявлення вразливостей у самому програмному коді.

В першому випадку, нейронні мережі можуть використовуватись для виявлення вразливостей у вхідних даних програмного продукту. Наприклад, у веб-додатку, який приймає дані від користувачів, нейронні мережі можуть бути використані для виявлення потенційно шкідливих введених даних, таких як SQL-ін’єкції, кросс-сайт скрипти та інші.

У другому випадку, нейронні мережі можуть використовуватись для виявлення вразливостей у самому програмному коді. В цьому випадку, нейронна мережа будується на основі великої кількості прикладів коду, які відповідають різним типам вразливостей у нових програмних продуктах.

Наприклад, нейронні мережі можуть використовуватись для виявлення SQL-ін’єкцій у програмному коді. Навчальний набір може містити приклади коді з SQL-запитами, які містять ін’єкційні фрагменти, а також приклади коду без ін’єкцій. Нейронна мережа навчається на цих прикладах і може виявляти ін’єкційні атаки у новому коді.

Одним з головних переваг використання нейронних мереж тут є те, шо вони можуть виявляти вразливості, які складно виявити іншими методами. Наприклад виявляти вразливості, шо пов’язані з певним контекстом або залежностями між різними фрагментами коду, а також виявляти вразливості, які не мають відповідної сигнатури в базі даних відомих вразливостей.

Однак, використання нейронних мереж у виявленні вразливостей в програмному забезпеченні має свої обмеження та недоліки. Перш за все, для навчання нейронних мереж потрібна велика кількість прикладів коду, які містять вразливості. Для того, щоб нейронна мережа була ефективною, вона повинна бути навчена на широкому спектрі вразливостей, що може бути часом трудомістким та вимагати великої кількості ресурсів. Крім того, нейронні мережі можуть виявляти вразливості з певним рівнем точності, але вони не можуть забезпечити 100% захист від усіх можливих атак, що спрямовані на їх навчання. Якщо зловмисник змінить навчальний набір даних для нейронної мережі, то вона може навчитись виявляти невірні вразливості або пропустити реальні вразливості. Також можливі атаки на саму нейронну мережу, такі як атаки з введенням гуму, що можуть спотворити її роботу.

Нейронні мережі можуть використовуватися для виявлення шкідливих програм, таких як віруси, троянські коні та різноманітні форми зловмисного програмного забезпечення. Шкідливі програми, такі як віруси, троянські програми, черви та шпигунські програми, можуть завдавати значних збитків користувачам та компаніям, які вони атакують. Виявлення цих програм є складним завданням, оскільки вони можуть бути масковані під легітимний код та мати високий рівень внутрішньої складності. У таких випадках використання нейронних мереж може допомогти знайти шкідливі програми з більшою точністю та швидкістю.

Перші спроби використання нейронних мереж для виявлення шкідливих програм почалися в 90-х роках минулого століття, коли були створені перші антивірусні програми з використанням нейронних мереж. Протягом останніх років, з введенням глибокого навчання та зростанням обчислювальних

можливостей, нейронні мережі стали ще більш ефективним інструментом для виявлення шкідливих програм.

Одним з найбільш ефективних підходів до виявлення шкідливих програм з використанням нейронних мереж є аналіз поведінки програми. У цьому підході нейронна мережа навчається виявляти ненормальну поведінку програми, таку як відкриття веб-сторінок без згоди користувача або передача даних на зовнішні сервери. Для цього можна використовувати різні види даних, наприклад, системні виклики, мережевий трафік та інші характеристики поведінки програми.

Інший підхід полягає у використанні нейронних мереж для аналізу коду програми на предмет наявності шкідливих функцій та вразливостей. В цьому випадку, нейронна мережа навчається виявляти певні шаблони коду, які можуть вказувати на наявність шкідливих функцій. Для цього можна використовувати різні формати вхідних даних, наприклад, текстовий код програми або виконуваний код в пам'яті.

Усі ці підходи, як і у випадку з виявленням вразливостей в програмному забезпеченні, вимагають великої кількості даних для тренування нейронних мереж. Для аналізу поведінки програми можуть використовуватись дані, зібрані від користувачів різних комп'ютерів, або створювати штучні датасети, що моделюють різні типи поведінки програм. Для аналізу коду програми можуть використовуватись навчальні датасети, складені з реальних шкідливих програм або створювати штучні датасети, що моделюють різні типи шкідливих функцій та вразливостей.

Одним з найбільш відомих прикладів використання нейронних мереж для виявлення шкідливих програм є система DeepLocker, яка була розроблена в IBM Research. Система використовує нейронні мережі для створення спеціальної програми, яка може запустити шкідливий код лише в певних умовах, наприклад, коли виявляється конкретна ціль атаки або коли виконується певна послідовність дій.

Нейронні мережі можуть бути використані для виявлення вторгнень у комп’ютерні системи. Нейронні мережі можуть аналізувати мережевий трафік та виявляти підозрілі дії, такі як відправлення незвичних запитів або спроби входу до системи з нових, невідомих IP-адрес.

Системи, які використовуються для виявлення вторгнень, можна класифікувати на два основних типи: системи виявлення вторгнень на основі правил (англ. rule-based intrusion detection systems, IDS) та системи виявлення вторгнень на основі машинного навчання (англ. machine learning-based intrusion detection systems, ML-based IDS)[36].

Системи виявлення вторгнень на основі правил зазвичай використовують експертну базу знань про типові атаки та відхилення від нормальної поведінки. Ці системи використовують правила, які описують характеристики атак та виявляють їх на підставі цих правил. Наприклад, якщо виявлено спробу доступу до системи з неправильними обліковими даними або надмірну кількість несправних запитів, система може спрацювати і сповістити про можливу атаку.

Системи виявлення вторгнень на основі машинного навчання використовують алгоритми машинного навчання для навчання моделей, які можуть виявляти вторгнення на основі вхідних даних. Ці системи використовують великий обсяг навчальних даних, щоб побудувати моделі, які виявляють аномалії або атаки. Можуть використовуватись різні типи алгоритмів машинного навчання, такі як нейронні мережі, метод опорних векторів, дерева рішень тощо.

Нейронні мережі можуть бути використані для класифікації даних у контексті інформаційної безпеки. Вони можуть фільтрувати спам та небажані листи на електронній пошті, класифікувати веб-сторінки на основі їх вмісту та визначати потенційно шкідливі та небезпечні сторінки. Цей підхід має так само великі переваги у порівнянні з традиційними методами класифікації даних в контексті інформаційної безпеки, оскільки контент який захований в даних

може бути дуже сильно видозмінений відносно шаблонів із якими порівнюють традиційні інструменти класифікації шкідливих даних.

Якісна класифікація та фільтрація даних, які приходять до користувачів інформаційних систем через відкриті шляхи комунікації такі як електронна пошта, можуть мінімізувати ризик, пов’язаний із людським фактором. Це дуже важливо в умовах сучасного бізнесу. Оскільки підприємства, які мають велику кількість співробітників, не можуть досконало моніторити дії кожного співробітника та гарантувати обізнаність та уважність кожного співробітника у сфері інформаційної безпеки.

Сучасні системи захисту інформації дедалі використовують все більш надійні методи аутентифікації та ідентифікації. Сучасні інструменти обмеження доступу можуть на вході обробляти біометричні дані, такі як розпізнавання обличчя, відбитків пальців, розпізнавання голосу тощо. Такі біометричні дані мають дуже велику кількість параметрів, які потрібно швидко і правильно обробити. Тому такі системи зчитування та порівняння біометрії використовують нейронні мережі. Нейронні мережі при отриманні даних біометрії можуть дуже швидко адаптуватися до чинників, які видозмінюють біометричну інформацію. Наприклад при розпізнаванні обличчя, алгоритми, які використовують нейронні мережі, можуть правильно розпізнати на людині аксесуари та зміни у зовнішності, такі як зміна зачіски, тощо. Тому в певному сенсі використання нейронних мереж відкрило можливість використання найскладніших типів біометрії на підприємствах, оскільки адміністраторам системи не потрібно буде кожен день змінювати біометричну базу даних відповідно до того, як співробітники мінімально змінюють свою зовнішність.

Також нейронні мережі використовуються в класичних процесах аутентифікації. Вони можуть ефективну збирати інформацію про звичайну поведінку користувачів системи при аутентифікації і подальшому використовувати цю інформацію для свого ж навчання. Таким чином система зможе відрізняти випадки зловмисної поведінки в процесі аутентифікації блокувати доступ.

Нейронні мережі можуть бути використані для прогнозування подій, які можуть вплинути на інформаційну безпеку. Нейронні мережі мають можливість прогнозувати атаки на мережі або визначати ризик витоку конфіденційної інформації. Вони можуть аналізувати історичні дані про атаки, порушення безпеки та інші фактори, щоб зробити прогнози стосовно можливих майбутніх загроз. Важливим аспектом використання нейронних мереж, який дуже корисний при прогнозуванні - це здатність виявляти нові, раніше невідомі загрози безпеці. Традиційні методи безпеки можуть бути обмежені в своїй ефективності через потребу в заздалегідь визначених правилах та сигнатурах загроз. У той час як нейронні мережі можуть адаптуватись до нових ситуацій та виявляти непередбачувані аномалії.

Криптографія - це наука про захист інформації шляхом шифрування та дешифрування даних. Нейронні мережі можуть бути використані для покращення безпеки криптографічних систем, а також для зламування криптографічних алгоритмів. Одним з основних способів використання нейронних мереж у криптографії є створення нових алгоритмів шифрування. Мережі можуть бути навчені використовувати складні математичні функції для шифрування даних. Вони можуть навчатися на великій кількості зразків шифрованого тексту та відкритого тексту, щоб встановити зв'язок між ними і створити алгоритм шифрування, який здатний ефективно захищати дані. З боку криптоаналізу нейронні мережі використовуються для аналізу великих обсягів даних і виявлення певних шаблонів, які можуть допомогти зламати криптографічний ключ. Наприклад, нейронні мережі можуть бути навчені для пошуку певних характеристик в шифрованому тексті, які пов'язані зі структурою ключа або залежностями між бітами. Ці характеристики можуть допомогти зламати криптографічний алгоритм і отримати доступ до зашифрованих даних.

Однак, варто зазначити, що використання нейронних мереж у криптографії є двобічною практикою. З одного боку, це може сприяти розробці більш ефективних криптографічних алгоритмів і забезпеченню більшої безпеки

даних. З іншого боку, це може становити загрозу для безпеки, оскільки зловмисники можуть використовувати нейронні мережі для зламування криптографічних систем і отримання незаконного доступу до конфіденційної інформації. Але в будь-якому випадку нейронні мережі дозволяють покращити аспект використання криптографічних методів захисту інформації.

Нейронні мережі використовуються для виявлення підозрілих дій користувачів в системах. Зокрема виявляють аномалії та незвичайну поведінку користувачів в системах, виявлення зловмисних дій, виявлення спроб несанкціонованого доступу та викрадення інформації, тощо. В таких системах захисту, які виконують аналіз поведінки користувачів, нейронні мережі тренуються на детально пропрацьованих шаблонах поведінки користувачів. В подальшому використанні ці системи аналізують поведінку кожного користувача, ідентифікують шкідливу поведінку користувачів і в подальшому результат кожної ітерації системи буде слугувати опорою для подальшої роботи системи.

Оцінка ризиків досить складний аналітичний процес, який потребує багато часу аби обробити та проаналізувати великий обсяг інформації пов’язаний із цільовою системою, її відповідність вимогам до захисту інформації, історію інцидентів та інші фактори, які впливають на ризики інформаційної безпеки. Використання нейронних мереж дозволяє автоматизувати цей процес та дає можливість створити системи аналізу ризиків, які можуть дати оцінку опираючись на інформацію про попередні інциденти, відомі вразливості функціоналу, яку використовує цільова система, конфігурація системи та інші фактори. Нейронні мережі можуть виявляти складні залежності та шаблони, які потребують набагато більше часу для аналізу у випадку використання інших методів.

Таким чином на даний момент вже існують основні вектори застосування нейронних мереж до інформаційної безпеки. Всі вони тим чи іншим чином пов’язані з обробкою величезного обсягу інформації про різні типи існуючих загроз. Окрім звичайного виявлення загроз, нейронні мережі мають здатність

використовувати отриману інформацію з кожного використання мережі в майбутньому, тим самим покращуючи якість та точність роботи мережі з кожною ітерацією.

В сучасному світі, де великі обсяги інформації зберігаються та передаються за допомогою комп'ютерних систем і мереж, проблеми інформаційної безпеки стають особливо актуальними. Швидкий розвиток технологій та зростання впливу цифрових платформ на всі сфери життя вимагають належного захисту конфіденційності, цілісності та доступності даних.

Зламання систем безпеки, хакерські атаки та поширення зловмисних програм становлять постійну загрозу для інформаційної безпеки. Нестача свідомості та навичок користувачів щодо захисту своїх даних додає до цього додаткові ризики. У такому контексті аналіз проблем інформаційної безпеки стає необхідним для розуміння вразливостей та розуміння шляхів використання нейронних мереж. Цей пункт досліджує основні виклики, з якими стикаються організації та індивідуальні користувачі у сфері інформаційної безпеки.

Злочинці здатні використовувати різноманітні кібератаки з використанням шкідливого програмного забезпечення, щоб отримати несанкціонований доступ до конфіденційної інформації. Існує дуже великий різновид шляхів, методів та інструментів за допомогою яких зловмисники можуть створювати програмне забезпечення. Таким чином кількість різних видів шкідливого програмного забезпечення з часом різко збільшується.

Традиційні алгоритми [37], які використовуються в засобах технічного захисту інформації для розпізнання шкідливого програмного забезпечення, такі як метод сигнатурного аналізу, циклічна перевірка контрольної суми, підрахунок файлової ентропії, використовують порівняльний аналіз, який представляє собою порівняння фрагментів вхідного досліджуваного зразка із правилами та фрагментами із бази знань. Основний виклик для сфери інформаційної безпеки становить саме велика різноманітність шкідливого програмного забезпечення. Постійне оновлення та підтримка бази знань систем

детектування шкідливого програмного забезпечення вимагає великих витрат, які через постійний розвиток шкідливого програмного забезпечення тільки збільшуються.

Сучасні пристрої, які ми використовуємо для різних цілей кожного дня, такі як побутові прилади, камери спостереження і навіть автомобілі, з розвитком інженерії отримали можливість підключення до мережі інтернет, що продукувало утворення інтернету речей. Згідно визначення Національного Інституту Стандартів та Технологій США [38], Інтернет Речей (IoT) -це мережа пристроїв, що містить апаратне, програмне забезпечення, прошивку та виконавчі механізми, які дозволяють пристроям з'єднуватися, взаємодіяти та вільно обмінюватися даними та інформацією. Багато підприємств використовують пристрої, які через наявність каналів комунікації можуть бути в зоні ризику кібератак, і як результат можуть бути виведені з ладу, що призведе до величезних збитків.

Фішинг є серйозною проблемою для багатьох підприємств індустріального сектору. Це метод атак соціальної інженерії, який полягає у шахрайському отриманні конфіденційної інформації, такої як паролі, фінансові дані або доступ до систем, шляхом впливу на людей і змушення їх виконувати певні дії[39].

Однією з основних форм фішингу є спам-електронні листи або повідомлення, які виглядають так, ніби вони відправлені відомими компаніями або особами. Ці повідомлення можуть містити посилання на підроблені веб-сайти, які дуже схожі на офіційні, але фактично створені зловмисниками. Коли співробітник підприємства переходить за таким посиланням і вводить свої облікові дані, зловмисник отримує доступ до цієї інформації і може використовувати її для шахрайства або несанкціонованого доступу до систем підприємства.

Атаки DoS і DDoS є серйозною проблемою для сучасних підприємств. Ці типи атак спрямовані на перевантаження ресурсів мережі або системи, змушуючи їх недоступними для законних користувачів.

DoS-атака (Denial of Service) – це атака, спрямована на вимкнення комп’ютера або мережі, що робить їх недоступними для користувачів. DoS-атаки досягають цього, переповнюючи ціль трафіком або надсилаючи їй інформацію, яка спричиняє збій. В обох випадках DoS-атака позбавляє законних користувачів (тобто співробітників, членів або власників облікових записів) сервісу або ресурсу, на який вони розраховували.

DDoS-атака (Distributed Denial of Service) – модифікація DoS атаки, в якій для відправки шкідливого трафіку використовується велика кількість машин для збільшення обсягу трафіку, який надсилається на цільову систему.

Атаки DoS і DDoS можуть призвести до значного зниження продуктивності підприємства. Коли ресурси системи або мережі зайняті обробкою великого обсягу шкідливого трафіку, легітимні користувачі можуть стикатися зі значними затримками або повністю втратити доступ до послуг. Якщо підприємство не здатне запобігти атакам DoS або DDoS, це може призвести до втрати довіри клієнтів. Користувачі, які не можуть отримати доступ до важливих послуг або сайту, можуть перейти до конкурентів, що негативно впливає на репутацію підприємства та його прибутковість. Також такі атаки призводять моментальних збитків, через потребу у відновленні системи підприємства після атаки.

1.4 Сучасні тенденції в розвитку Штучного Інтелекту для кібербезпеки

Слідкувати за тенденціями у сфері ШІ важливо, щоб бути в курсі останніх подій та бути на передовій технологічних інновацій. Ці знання дозволяють досліджувати нові можливості, адаптуватися до нових викликів і активно сприяти розвитку індустрії ШІ [40].

МЕНЮ 80% керівників використовують технологію ШІ для розробки стратегії та прийняття бізнес-рішень. Очікується, що принаймні кожна десята компанія інвестуватиме у створення цифрового контенту з використанням ШІ.

Добра поінформованість також покращує здатність брати участь у

змістовних дискусіях, долучатися до проектів і зберігати вигідну позицію в умовах, що швидко змінюються. Зрештою, оновлюючись, ентузіасти зможуть використовувати весь потенціал штучного інтелекту і приймати впевнені рішення у своєму професійному та особистому житті.

Штучний інтелект лідирує в тому, щоб зробити цифровий світ безпечнішим. Ось кілька способів, як це зробити:

- Використання передових алгоритмів У 2024 році ШІ буде використовувати передові алгоритми для глибокого занурення в цифрове середовище і безперервного сканування на наявність потенційних загроз.
- Реагування в режимі реального часу: ШІ миттєво виявляє загрози і миттєво реагує на них. Реагування в режимі реального часу мінімізує потребу хакерів у використанні вразливостей.
- Поведінковий аналіз для більшої точності: ШІ не обмежується розпізнаванням відомих загроз. Інтегруючи поведінковий аналіз, він дізнається, що є "нормальним" для кожного користувача; ШІ може виявляти відхилення від стандартної поведінки і показувати ознаки потенційних проблем з безпекою до того, як вони стануть повномасштабними інцидентами.
- Швидке реагування завдяки виявленню аномалій Виявлення аномалій - це як цілодобове чергування охоронця: ШІ виявляє порушення і швидко реагує, щоб точно ідентифікувати і нейтралізувати потенційні загрози безпеці.
- Мінімізація вікон вразливості: ШІ не придушує кіберзагрози. Зменшуючи вікна вразливості - вікна, в яких системи піддаються потенційним атакам, - ШІ гарантує, що цифрові фортеці захищені і завжди випереджають кіберсупротивників.
- Сприяння цілеспрямованому реагуванню: не існує універсального рішення; ШІ адаптує свою реакцію на основі конкретних загроз, з якими він стикається. Цільовий підхід означає менше супутніх збитків і більш точне реагування на інциденти безпеки.
- Захисний ефект штучного інтелекту: Коли ШІ виступає в ролі цифрового захисника, кібербезпека стає проактивною, а не реактивною.

Йдеться не лише про боротьбу із загрозами, а й про передбачення, запобігання та випередження у постійній війні з кібернетичними супротивниками.

У 2024 році архітектура нульової довіри з використанням штучного інтелекту розвиватиметься з удосконаленнями, які підвищать її ефективність у сфері кібербезпеки. Цей підхід використовуватиме принцип "не довіряй нікому, перевіряй усе" і використовуватиме ШІ для подальшого вдосконалення процесу безперервного оцінювання.

Пристосування засобів контролю доступу до мінливих ризиків стане більш досконалим, що дасть змогу здійснювати постійний і ретельний моніторинг облікових даних і діяльності користувачів. Виявлення аномалій за допомогою штучного інтелекту дозволить Zero Trust виявляти незвичайні патерни і більш точно реагувати на них, зміцнюючи таким чином свою систему безпеки.

Комісія з цінних паперів і бірж (SEC) працює над тим, щоб задовольнити довгострокові вимоги Zero Trust, встановлені Адміністративно-бюджетним управлінням. Федеральні агентства повинні досягти цілей безпеки без довіри до кінця 2024 фінансового року. Для цього відомства повинні призначити керівника стратегії нульової довіри і виконати 19 завдань.

З огляду на те, що багато факторів відіграють роль, коли ШІ оцінює поведінку користувача і місцезнаходження пристрою, такий підхід до безпеки буде необхідний для забезпечення індивідуальних заходів для конкретних ситуацій. Штучний інтелект у резервному копіюванні та відновленні даних

Це стане стандартною практикою до 2024 року і змінить підхід організацій до безпеки. У випадку з Кіотським університетом 77 терабайт дослідницьких даних було втрачено через погану систему резервного копіювання.

Помилка сталася через те, що останнє завдання резервного копіювання негайно перезаписало попередню резервну копію, що зробило її непридатною для використання, коли потрібно було відновити дані. Поява продуктивних інструментів зі штучним інтелектом вносить зміни в процес аварійного відновлення. Вони роблять процедури відновлення більш ефективними та

надійними, ніж традиційні методи.

Це дозволяє організаціям передбачити значне підвищення стійкості даних і забезпечити більш надійний захист від потенційної втрати або пошкодження. Роль штучного інтелекту також поширюється на оптимізацію робочих процесів відновлення.

Таке швидке та ефективне відновлення має вирішальне значення для підтримки безперервності бізнесу та пом'якшення потенційного впливу кібератак.

Ворожий ШІ, розроблений для того, щоб перевершити інші системи ШІ, стає вектором загрози.

Щоб протистояти ворожому ШІ, організації повинні стратегічно інвестувати в стійкі системи. Для підвищення стійкості потрібні жорсткі методи навчання моделей. Механізми безперервного моніторингу відіграють вирішальну роль у виявленні та пом'якшенні наслідків атак.

Боротьба з ворожим ШІ вимагає співпраці в рамках спільноти кібербезпеки. Обмін інформацією, тактикою і стратегіями захисту має важливе значення для того, щоб випереджати нові загрози. Єдиний фронт сприяє адаптивності і забезпечує сильніший захист.

До 2024 року операції з кібербезпеки зазнають трансформації, в якій на перший план вийдуть штучний інтелект і людський досвід: Інструменти на основі штучного інтелекту розширять можливості фахівців з кібербезпеки, покращуючи їхні можливості прийняття рішень і реагування.

Ця інтеграція має на меті досягти балансу, дозволивши штучному інтелекту ефективно виконувати повсякденні завдання, а аналітикам зосередитися на поглибленому аналізі та стратегічному плануванні. Ця синергія створить надійну та адаптивну робочу силу з кібербезпеки, яка зможе ефективно протистояти кіберзагрозам.

У 2022 році майже половина компаній стануть жертвами кібератак за участю третіх осіб. Того ж року було здійснено понад 112 мільйонів атак на системи Інтернету речей. Ось як технології штучного інтелекту, що зберігають

конфіденційність, вплинуть на кібербезпеку в 2024 році:

Передові технології: Зважаючи на проблеми конфіденційності, організації використовують передові технології, такі як спільне навчання та гомоморфне шифрування.

Безкомпромісні інсайти: Ці технології дозволяють організаціям отримувати цінну інформацію з даних, не порушуючи при цьому приватність користувачів.

Відповідність нормативним вимогам: ШІ, що зберігає конфіденційність, безперешкодно працює з регуляторними вимогами, що постійно змінюються, і забезпечує міцну основу для дотримання нормативних вимог.

Побудова довіри: Цей підхід зміцнює довіру між користувачами та зацікавленими сторонами, наголошуючи на відповідальному поводженні з конфіденційною інформацією.

Досягнення балансу: ШІ, що зберігає конфіденційність і забезпечує баланс між ефективними заходами кібербезпеки та повагою до права на приватність, є наріжним каменем етичного та безпечного управління даними.

Регуляторні органи надають великого значення прозорості та підзвітності. Необхідність відкритості алгоритмів штучного інтелекту стає все більш важливою для дотримання нормативних вимог. Наочні моделі ШІ важливі, оскільки організаціям потрібно показати, як приймаються рішення з використанням ШІ. Ці моделі забезпечать чітке розуміння процесу прийняття рішень і полегшать перевірку відповідності.

За оцінками, до 2030 року 30% завдань буде автоматизовано за допомогою технології штучного інтелекту. Для підготовки до нової ери навчання з кібербезпеки, в якій ШІ відіграватиме активну роль, вам потрібно знати:

- Реалістичні навчальні сценарії: Симуляційні платформи на основі ШІ створюють реалістичні навчальні сценарії, які відображають складність динамічних загроз.
- Адаптація до нових загроз: Навчальні модулі зі штучним інтелектом адаптуються до загроз. Це дозволяє фахівцям з кібербезпеки розвивати свої

навички, постійно стикаючись з новітніми викликами.

- Покращений розвиток навичок: Впровадження ШІ покращує розвиток навичок і забезпечує практичний і захоплюючий досвід. Фахівці можуть відточувати свої навички в контрольованому середовищі, перш ніж зіткнутися з реальними кіберзагрозами.
- Прискорене навчання: Навчання з використанням штучного інтелекту прискорює процес навчання для новачків у сфері кібербезпеки. Ці модулі адаптуються та налаштовуються, що дозволяє фахівцям швидко розібратися в складнощах цієї галузі.

Висновки до розділу 1

В першому розділі було представлено історію розвитку штучного інтелекту та основні можливості застосування його в рамках кібербезпеки. Таким чином штучний інтелект вже має потенційне місце для використання в кібербезпеці. Він має перевагу перед старими рішеннями з кібербезпеки, оскільки має можливість зекономити витрати на модифікування систем кібербезпеки відповідно до нових загроз, а також надає можливість виявляти раніше невідомі загрози. Стрімке збільшення попиту навикористання штучного інтелекту в різних сферах, а також збільшення самого обсягу програм, додатків та технологій з використанням штучного інтелекту створює величезний виклик для сфери кібербезпеки. В той же самий час великі обсяги інформації та збільшення кількості нових видів атак ускладнюють процеси виявлення загроз за допомогою технічних засобів захисту інформації побудованих на традиційних методах аналізу загроз. В наступних розділах буде розглянуто, які функції штучного інтелекту в кібербезпеці вже існують, а також, які існують готові рішення для кібербезпеки з використанням штучного інтелекту на даний момент.

Розділ 2 ЕФЕКТИВНІСТЬ ЗАСТОСУВАННЯ ШТУЧНОГО ІНТЕЛЕКТУ В КІБЕРБЕЗПЕЦІ

2.1 Автоматизація процесів виявлення та аналізу кіберзагроз за допомогою ШІ

Штучний інтелект може автоматично збирати, обробляти та аналізувати дані з різноманітних джерел, пов’язаних з інформаційною безпекою: логи подій, результати аудиту, сповіщення систем безпеки, звіти про інциденти тощо.

Алгоритми машинного навчання та нейромережі ефективно виявляють у таких даних патерни, кореляції та аномалії, що можуть сигналізувати про вразливості системи, спроби кібератак та інші проблеми. ШІ здійснює такий аналіз в десятки разів швидше і точніше за людину.

Засоби ШІ можуть ефективно фіксувати спроби втручання у систему ззовні. Для цього використовуються два основних підходи: розпізнавання сигнатур та евристичний аналіз.

Метод розпізнавання сигнатур передбачає пошук вже відомих алгоритмів зламу: ШІ порівнює вхідні дані із базою відомих сигнатур атак (вірусні коди, експлойти тощо), аби миттєво виявляти та зупиняти атаки за поширеними сценаріями. Метод евристичного аналізу передбачає вивчення поведінки системи та користувачів, аби виявляти відхилення від норми, що можуть сигналізувати про атаку. Важливо, що алгоритми ШІ можуть здійснювати таку оцінку безперервно та в режимі реального часу, та реагувати миттєво, навіть без участі людини.

Напрямок моніторингу і аналізу трафіку напряму пов’язаний з попереднім. ШІ допомагає слідкувати за потоками даних у системі, аналізуючи трафік між пристроями, додатками та сервісами. Подібним чином він слідкує за тим, як з системою взаємодіють користувачі: які вони авторизуються, які файли завантажують на сервери, як обмінюються повідомленнями тощо. ШІ може використовувати методи аналізу мережного трафіку та поведінкового аналізу

для виявлення аномалій, які можуть вказувати на наявність атаки чи компрометації даних. Він також може також адаптуватися до змін, навчаючись на нових даних та оновлюючи свої моделі.

Хибні сповіщення шкодять кіберзахисту, адже погано впливають на сприйняття небезпеки та призводять до надмірних витрат і без того обмежених ресурсів безпекового сектору.

Аби система сповіщень про загрози кібербезпеки працювала коректно, необхідно по-справжньому майстерне налаштування алгоритмів безпеки та ручний моніторинг. Штучний інтелект в інформаційній безпеці бізнесу допомагає знизити кількість хибних сповіщень, спричинених неправильною оцінкою звичайних подій та нешкідливих епізодів у роботі системи.

Використання методів статистичного аналізу, алгоритмів ML та нейромереж дозволяє машинам оцінювати загрози не гірше за людину, враховувати в оцінці зворотний зв’язок від фахівців та контекст епізоду.

Засоби ШІ можуть допомогти з оцінкою та оптимізацію при масштабному оновленні IT-інфраструктури в компанії. Наприклад, при встановленні нової системи над новим локальним середовищем, при переході в хмару, при впровадженні нових технологій або інтеграції різних систем тощо. Алгоритми ШІ спрощують аналіз конфігурації та налаштування, а також тестування для перевірки сумісності, продуктивності та безпеки системи.

Алгоритми тестування ШІ дозволяють виявити більшість конфліктів, помилок та вразливостей, які можуть виникнути при модернізації або зміні робочого середовища. Досягнути подібних результатів лише мануальним тестуванням практично неможливо.

Алгоритми ШІ можуть бути корисними у класифікації та кластеризації даних системи під різноманітні вимоги. Зокрема – для дотримання вимог законодавства у частині інформаційної безпеки, для подальшої побудови профілів атак та вразливостей, для аналізу даних в контексті епізодів кібератак, а також для подальшого прогнозування та формування стратегії кіберзахисту.

2.2 Використання Штучного Інтелекту для прогнозування та управління кіберризиками

CPS та IoT виробляють величезну кількість даних, і для аналізу таких великих даних потрібні передові аналітичні інструменти. Щоб усунути шум і суперечливість даних, нам майже напевно знадобляться аналітичні інструменти зі штучним інтелектом [41]. З точки зору потоків даних, IoT було описано як революційне вдосконалення технології, яке змінює традиційне життя на стиль життя з високими технологіями. З іншого боку, архітектури CPS представляють дуже широку концепцію. Система повинна інтегрувати ці різноманітні концепції у когнітивний стан для аналітики великих даних і статистичного машинного навчання для прогнозування кіберризиків. Але розробка систем великих даних для периферійних обчислювальних середовищ є складною.

Одним із найактуальніших моментів для CPS є, можливо, безпека, як електронні, так і фізичні, що пов’язує фізичні та кіберсистеми. Така безпека вимагає гарантії інформації та захисту даних, що передаються з фізичних та електронних доменів і засобів зберігання. Крім того, керування активами та контроль доступу потрібні для надання або відхилення запитів до служб обробки інформації, особливо через те, що CPS буде взаємодіяти з нетехнічними користувачами та оскільки можливий вплив через адміністративні межі. Потрібні методи усунення нових вразливостей, викликаних проблемами життєвого циклу, включаючи зменшення джерел виробництва та оновлення активів. Вони включають підходи до проектування системної динаміки в різних часових масштабах, наприклад, архітектури з незначним часом та управління динамікою конструкції.

Крім того, CPS вимагає боротьби з підробками та управління ризиками ланцюга постачання для протидії зловмисним компонентам ланцюга постачання, які були змінені в порівнянні з їх оригінальним дизайном, щоб спричинити збій або несанкціоноване функціонування. Поряд зі стандартизацією дизайну та процесу, гіперзв’язок у цифровому ланцюжку

поставок також потребує підтримки. Передбачається, що обмеження доступу до вихідного коду для критично важливого та кваліфікованого персоналу може забезпечити гарантію програмного забезпечення та безпеки додатків і може бути необхідним для усунення навмисного впровадження AWS та вразливостей у CPS [42].

Заходи безпеки повинні включати криміналістику, прогнозування та плани відновлення для аналізу кібератак і для координації з іншими CPS та організаціями, які ідентифікують зовнішні вектори кібератак. Щоб вирішити цю проблему, внутрішній мережевий процес відстеження може допомогти у виявленні або запобіганні наявності слабких місць у засобах контролю безпеки логістики. Для підтримки цього необхідний процес розробки системи захисту від шкідливих зловмисників і втручання, щоб запобігти використанню вразливостей CPS, виявлених під час атак зворотного проектування.

Таблиця 2.1 Таксономія зон фокусування (AoF) для механізму когнітивного зворотного зв'язку в прогнозній аналітиці кіберризиків

Зони фокусування (AoF) підкреслили необхідність конфіденційності в механізмі зворотного зв’язку для звітів про кібератаки та спільних базах даних у аналізі ризиків CPS. У розділі, систематичний аналіз застосовано до кожної фокусної області, щоб визначити її збіг з літературою про штучний інтелект у прогнозній аналітиці кіберризиків CPS.

«Сервітизація» — це перехід від продажу фізичних продуктів до продажу поточних послуг, які ці продукти надають, або поточних послуг, які підтримують роботу продуктів. У контексті штучного інтелекту в аналітиці ризиків CPS ці послуги включають прогнозне технічне обслуговування, прогнозування відмови машини та автоматичну діагностику несправностей. Наприклад, інтелектуальні алгоритми машинного навчання отримують інформацію від промислових датчиків і платформ Інтернету речей, щоб автоматично діагностувати несправності та оцінювати залишковий термін служби обладнання.

В обґрунтованій теорії для розробки таксономій застосовується метод обґрунтованої теорії (GT), щоб згрупувати вимоги до штучного інтелекту для аналізу ризиків CPS для «сервітизації» у виробництві. Аналіз обґрунтованої теорії вбудовано в концептуальну діаграму, що представляє каскадний ієрархічний процес через сфери фокусування безпеки CPS (рис.2).

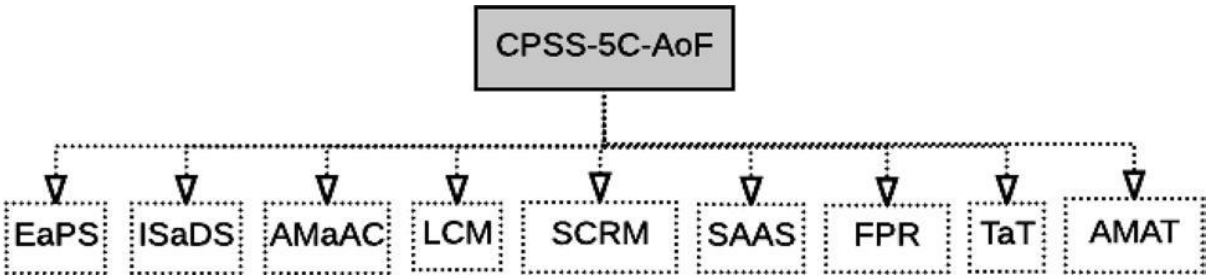

Рис. 2 Безпека CPS у сферах фокусування в п'ять рівнів CPS

Це інструмент для візуалізації зон фокусування, отриманих в результаті аналізу. Області фокусування (рис.2). Останні класифікуються на п'яти рівнях штучного інтелекту в CPS (див. табл.2.1).

Таблиця 2.1

Таксономія областей фокусування для штучного інтелекту для аналізу ризиків CPS

Безпека CPS	CPSS
Сфери фокусування	AoF
Архітектура 5C	5C
Електронна та фізична безпека	EaPS
Інформаційне забезпечення та безпека даних	ІСАДС
Управління активами та контроль доступу	AMaAC
Життєвий цикл і захист від підробок	LCM
Зменшення джерел виробництва, брак матеріалів та управління ризиками ланцюга поставок	SCRM
Гарантія програмного забезпечення та безпека програм	SAAS

Криміналістика, прогнозування та плани відновлення	FRP
Трек і слід	TaT
Захист від зловмисників і втручання	AMAT

EaPS - електронна та фізична безпека для штучних технологій. Для цього потрібні рішення для збору та зберігання даних у реальному часі для eets машин, що забезпечує адаптивний аналіз і одноранговий моніторинг.

ISaDS - забезпечення інформації та безпека даних для штучних технологій. Це має підтримуватися автономними когнітивними рішеннями, алгоритмами машинного навчання та високопродуктивними обчисленнями або аналізом даних, що підтримує швидкий обмін інформацією про кібератаки та звітування через спільні ресурси бази даних.

AMaAC - управління активами та контроль доступу для аналітики кіберризиків. У динамічній аналітиці кіберризиків це вимагає, щоб машини еволюціонували до CPS.

SCRM - життєвий цикл і захист від підробок для штучного інтелекту для аналізу кіберризиків. Для цього потрібні інтерфейси «людина-машина», для самосвідомих машин і прогнозування компонентів і управління працездатністю.

LCM - зменшення джерел виробництва, дефіцит матеріалів і управління ризиками в ланцюзі поставок. Це потрібно для встановлення пріоритетів та оптимізації рішень із самооптимізованими виробничими системами, що підтримується комп’ютерною візуалізацією планування виробництва, такою як інтеграція систем SCADA з віртуальною реальністю для розробки системи підтримки прийняття рішень.

SAAS - гарантія програмного забезпечення та безпека додатків для штучного пізнання. Для цього потрібна платформа великих даних для моніторингу стану датчиків. Такі платформи можуть створювати складні моделі, такі як дизайн кіберміста з використанням структурованого зв'язку для

мобільних CPS, міждоменне наскрізне спілкування між об’єктами та технології хмарних обчислень.

FPR - криміналістика, прогнозування та плани відновлення для штучного пізнання. Це потрібно інформувати за допомогою ключових показників ефективності.

TaT - відстеження в аналітиці кіберризиків. Механізми зворотного зв’язку та контролю необхідні для забезпечення наглядового контролю дій, щоб уникнути або надати необхідний доступ або розробити стійку систему контролю.

AMAT - захист від зловмисників і втручання. Цьому сприяли б архітектури з незначною активацією часу та контроль динаміки конструкції.

Вимоги до ШІ для CPS у виробництві та «сервітизації» представлені в класифікації зі скороченнями (табл.2.2.1), які підтримують надійну інтеграцію штучного інтелекту для аналізу кіберризиків. Таксономія представляє вимоги до ШІ, визначені в літературі про прогнозну аналітику кіберризиків, де компоненти та з’єднувачі ШІ обслуговують всю систему під час виконання.

Таксономія вимог для штучного інтелекту для CPS у виробництві та «сервітизації», забезпечує цілісне розуміння вимог до інтеграції когнітивного CPS у аналітику кіберризиків із динамічними даними в реальному часі від виробництва та «сервітизації». Групування вимог використовується для аналізу необхідних програм і технологій, а також для створення каскадної архітектури для інтеграції штучного інтелекту для CPS. Ця тема була визначена як імперативна в інженерній літературі, для оцінки впливу кіберризиків IoT.

2.3 Переваги та обмеження використання Штучного Інтелекту в кібербезпеці

Загальний вплив штучного інтелекту на кібербезпеку є як позитивним, так і негативним. Атаки на компанії стають дедалі небезпечнішими. Зловмисники намагаються збільшити свої знання, щоб знайти слабкі місця в

технологіях кібербезпеки. Автоматизація в результаті алгоритмів машинного навчання гарантує, що зловмисники не можуть використовувати ті самі способи для атаки на системи зі штучним інтелектом. Технологія продемонструвала, що алгоритми машинного навчання краще забезпечують безпеку, ніж люди. Інтеграція штучного інтелекту в кібербезпеку гарантує уникнення помилок. Цей фактор входить до числа переваг штучного інтелекту для кібербезпеки, про які йдеться нижче.

Різні технології штучного інтелекту відіграють різну роль у забезпеченні кібербезпеки. Технології продовжують досліджувати, щоб забезпечити максимальну ефективність уникнення атак. Організації по всьому світу мають конфіденційну інформацію, тому технології мають забезпечити захист, щоб ніхто не мав доступу до цієї інформації. Майбутнє також передбачає широке впровадження штучного інтелекту. Цей фактор означатиме, що штучний інтелект буде високорозвиненим для забезпечення максимальної безпеки в організаціях. Наявність систем, які могли б захистити себе та виявити будь-які спроби, є одним із бачень більшості компаній. Аспект безпеки — це мрія, за досягнення якої борються дослідники та ІТ-компанії. Першим ключовим атрибутом штучного інтелекту, який значною мірою вбудований у системи, є навчання на їхньому досвіді. Це одна з основних характеристик ШІ в цілому. Було доведено, що системи можуть навчатися з різних аспектів, які зробили цю технологію надзвичайно актуальною для кібербезпеки. ШІ розглядається як технологія «прийди і рятуй» у сфері кібербезпеки. Навчання на досвіді є атрибутом алгоритмів штучного інтелекту, де системи можуть навчатися на факторах, які були раніше. Алгоритми використовувалися в технологіях і алгоритмах кібербезпеки, щоб гарантувати, що помилка не повториться. Таким чином, атаки вбудовані в систему, де алгоритм штучного інтелекту буде виявляти атаку та навчатися.

Технологія ШІ є однією з найдосконаліших технологій у сучасному світі. Технологія вдосконалила все, що машина працює і чого хоче досягти. Людина відчуває невгамовну потребу створювати гаджети, які можуть бездоганно

виконувати обчислення та виконувати дії без обмежень. Технологія штучного інтелекту є одним із найкращих винаходів людини, навіть за межами нашого розуміння. Кожна організація, яка впроваджує технологію, гарантує, що вони покращили свої послуги та ефективність. Технологія штучного інтелекту також сприяла тому, що вона допомогла зменшити кількість кіберзлочинів, з якими ми стикаємося, оскільки це одна з проблем сучасного світу [43]. Технологія штучного інтелекту підтвердила, що ці дії виявляють та зменшують проблеми. Технологія штучного інтелекту забезпечила швидше виявлення збоїв у системі, оскільки їх кадри моніторингу набагато більші, ніж у людини. Цей аспект різко впливає на забезпечення відсутності злочинів або проникнення в систему неавторизованих осіб. І тому технології сприяли відмінній технологічній безпеці в сучасному світі. Моніторинг трафіку в реальному часі дозволяє технології штучного інтелекту ідентифікувати будь-яку активність без правильних вимірювань або протоколів і реагувати на дії. Цей фактор дозволяє технології впоратися з ситуацією, поки не пізно. На цьому етапі система буде безпечною та не пошкодженою, допомагаючи організації покращити протоколи безпеки та захистити свої дані та інформацію.

Що стосується заходів безпеки даних і протоколів, технологія ШІ була однією з найкращих технологій у забезпеченні їх покращення. Дані мають вирішальне значення для бізнес-організацій, тому їх необхідно захищати [44]. За допомогою кількох протоколів шифрування даних система може сприяти чудовому шифруванню та гарантувати безпеку залучених даних. Великий протокол значно впливає на технологію в секторі технологій кібербезпеки.

Технологія штучного інтелекту також призвела до безробіття на деяких посадах у сфері кібербезпеки, які вона замінила. Комп’ютер має кращу ефективність у всьому, що він робить, що робить його пріоритетом для людей, які мають навички в цій галузі. Введення технології в бізнес, вплинуло на працевлаштування спеціалістів з кібербезпеки, оскільки вони мали менший внесок в організацію, оскільки технологія штучного інтелекту робила все добре та ефективно. Це також призвело до того, що організація скоротила рівень

обслуговування та перевірок системи. Те, як технологія штучного інтелекту захищає протоколи безпеки, неймовірно ефективна порівняно з тим, де це робить окрема особа. Організації, які використовують цю технологію, гарантовано захищають свої дані, оскільки технологія в їхній системі забезпечує більшу ефективність, оскільки вона продовжує розуміти свою систему та операції.

Включення систем навчання штучного інтелекту в кібербезпеку допомагає запобігти атакам на систему. Система, заснована на навчанні, вивчає дії зловмисників і налаштовується для захисту інформації. Цей фактор унеможливлює доступ зловмисників до даних. Наявність системи, яка постійно адаптується та навчається, є одним із атрибутів, які зробили технологію дуже ефективною. ШІ зміг уникнути кібератак за допомогою підходів, розглянутих нижче. Різні методи забезпечують ефективність ШІ в кібербезпеці.

Подальший вплив штучного інтелекту на кібербезпеку відбувається через методи, засновані на сигнатурах. Розуміння кодів підпису було критично важливим атрибутом технології ШІ в кібербезпеці. Метод полягає в тому, що штучний інтелект виявляє кібератаки та шкідливі програми за допомогою доступних кодів. Ці коди в шкідливих програмах або атаках виявляються за допомогою алгоритму ШІ. Таким чином, зіставлення сигнатур з останніх атак або бази даних дає команді з кібербезпеки перевагу в припиненні атаки. Сигнатури необхідно швидко порівняти, щоб виявити атаку. Таким чином, тип атаки, який розуміється, забезпечує час і ресурси, необхідні для припинення атаки. До того, як технологія ШІ вплинула на кібербезпеку, ці виявлення проводилися багато часу, що призводило до масових збоїв і втрат.

База даних, згадана вище, де зберігаються сигнатури шкідливих програм, називається чорним списком. Система виявляє атаку шляхом порівняння доступних підписів у чорних списках із відомими підписами, виловленими під час атаки. Сигнатури іноді називають шаблонами, присутніми в атаці, і це можна назвати іншою формою машинного навчання. Хоча цей метод протягом багатьох років довів свою ефективність, у разі нової атаки він був марним.

Техніка не вдається, оскільки в базі даних немає запису про атаку. Однак ця технологія виявилася дуже ефективною та зупинила багато атак протягом багатьох років. Було помічено, що вищевказану техніку можна уникнути за допомогою спеціальних методів. Ключовим прикладом є те, як зловмисники зрозуміли різні способи уникнення атак шляхом зміни своїх шаблонів. Хакери розуміють аспекти, сприйнятливі до ШІ, і повністю змінюють ці значення. Зміна шаблонів гарантує, що вони зможуть отримати доступ до даних та інформації до того, як їх виявлять. Як згадувалося вище, метод показав величезний вплив на кібербезпеку. Дослідження показують, що більшість атак вдалося зупинити та уникнути за допомогою цього методу.

Як згадувалося вище, машинне навчання істотно вплинуло на кібербезпеку. Люди завжди роблять помилки, аналізуючи дані чи інформацію. Значною перевагою технологій ШІ є те, що вони системні. Перевага системи полягає в уникненні помилок або відсутності деталей атак. Використання ШІ для аналізу журналів і мережевих пакетів забезпечило швидке виявлення атак. Технологія ШІ виявляє системи, аналізує доступні записи та виявляє журнали, включені в систему. Цей фактор гарантує, що системні адміністратори можуть змінювати інформацію, до якої мають доступ, щоб уникнути подальших втрат. Цей фактор привів до аналогії, згідно з якою штучний інтелект тісно замінює людей-аналітиків.

Головна перевага штучного інтелекту в кібербезпеці – це здатність аналізувати величезні дані. Обширні дані завжди втомлюють досліджувати, будучи людиною-аналітиком. Цей атрибут був суттєво змінений після появи інформації про технології ШІ. ШІ може аналізувати великі фрагменти інформації і не робити помилок. Люди-аналітики також ефективні у виявленні, оскільки вони можуть керувати технологією ШІ. Зусилля між системами та аналітиками гарантують, що всі наявні дані були проаналізовані та порівняні. Цей фактор довів свою ефективність у припиненні атак. Перш ніж запобігти атакам і захистити доступні дані, ідентифікація зловмисного програмного забезпечення завжди є першим кроком. Таким чином, класифікація та

кластеризація є чудовими атрибутами систем машинного навчання. Вони порівнюють наявну інформацію та те, як вона має бути в журналах. Цей фактор забезпечує виявлення помилок у системі. Регулярні записи порівнюються з поточними, щоб визначити заражені журнали. Після виявлення атаки вживаються необхідні кроки, щоб забезпечити її припинення. Кластеризація передбачає групування доступних записів або інформації з системи та виявлення аномалій. Обидва ці методи, які використовуються в машинному навчанні, виявилися ефективними, оскільки це неможливо для людей.

Мережеві атаки є однією з найпоширеніших форм агресії в кібербезпеці. Рейди проводяться через мережі, які використовують організації чи компанії. Завжди важливо виявляти атаки через мережі. Цей фактор дає перевагу системі зупиняти атаку з інтернету. Завдяки штучному інтелекту цей атрибут став дуже простим. Мережеві брандмауери з вбудованою технологією ШІ також виявилися дуже ефективними. Доступ до мережі було дуже ускладнено без належної авторизації. Припинення атак з Інтернету є першим кроком у захисті доступної інформації. Таким чином, цей підхід виявився дуже ефективним у запобіганні майбутнім атакам. Наведені вище вказівки також вбудовані в мережі для забезпечення максимальної безпеки. Основною ключовою характеристикою та перевагою систем виявлення мережевих вторгнень є наявність у них п’яти елементів, які забезпечують повну безпеку таких мереж. Перший ключовий елемент полягає в тому, як системи ШІ отримують великі обсяги інформації з мережі. Цього фактора можна досягти завдяки здатності системи ШІ аналізувати великі обсяги даних. Усі фактори допомагають забезпечити безпеку мережі. Припинення атаки з мережі дає організації більше шансів захистити інформацію. За допомогою доступних методів штучного інтелекту можна уникати всіх способів зламу мережі.

Наведені вище дискусії демонструють, що ШІ вплинув на кібербезпеку у великих масштабах. Спосіб, згаданий вище, полягає у впливі штучного інтелекту на кібербезпеку на рівні мережі. Системи навчаються різним формам уникнення будь-яких можливих атак, щоб забезпечити безкомпромісність

мережі. Цей фактор навчання штучного інтелекту також відіграв значну роль у забезпеченні безпеки мережі. Цей фактор та багато інших збільшили переваги штучного інтелекту для ШІ, що впливає на кібербезпеку.

Управління вразливістю є атрибутом машин зі штучним інтелектом, які керують можливими вразливими місцями, які організації можуть мати у своїх системах. Дослідження показують, що у 2019 році було зареєстровано близько 20 362 уразливостей. Порівняно з 2018 роком кількість збільшилась на 18%. Цей фактор демонструє, що організації продовжують щодня стикатися з загрозами. Управління цими вразливими місцями стає виснажливим для присутнього персоналу. Цей фактор вимагав включення систем штучного інтелекту для керування зареєстрованими експозиціями. Цей фактор ускладнює доступ хакерів до систем. Таким чином, управління вразливістю є однією з переваг впливу ШІ на кібербезпеку. Згідно з дослідженням IBM про динаміку ринку штучного інтелекту в кібербезпеці, яке враховує всі виявлені вразливості, витрати на кіберпростір у всьому світі збільшуються, незважаючи на пандемію COVID-19 (Табл. 2.3.1).

Таблиця 2.3.1

Оцінка штучного інтелекту в прогнозуванні ринку кібербезпеки

Ринок	Цінність штучного інтелекту на ринку кібербезпеки
Розмір ринку 2018	9,8 млрд доларів США
Розмір ринку 2021	14,9 млрд доларів США
Розмір ринку 2025	36,6 млрд доларів США
Розмір ринку 2030	133,8 млрд доларів США

Більшість хакерів використовували повільну реакцію доступного керування вразливістю. Системи штучного інтелекту, які керують базою даних вразливостей, гарантують, що спроби атак повідомляються в режимі реального часу, таким чином забезпечуючи безпечніші системи.

Іншим важливим аспектом є те, як алгоритми машинного навчання виявляють аномалії облікових записів користувачів. Цей фактор гарантує, що системи захищені, якщо користувач у системі виявиться загрозою. Аспект управління вразливістю за допомогою систем штучного інтелекту забезпечив більшу безпеку серверів і безпечнішу інформацію, що зберігається на цих машинах.

Кібербезпека передбачає захист даних від будь-яких атак. Використання штучного інтелекту, як згадувалося вище, гарантує, що ці способи є більш ефективними та безпечними. Центри обробки даних є одним із найбільш критичних аспектів, які вимагають кібербезпеки. Виявлено, що головна перевага ШІ полягає в тому, що процеси, включені в ці центри, були автоматизовані. Енергоспоживання, використання пропускної здатності та температури є життєво важливими аспектами, які значно контролюються в центрах обробки даних. Оскільки люди іноді помиляються, використання ШІ для керування такими галопами забезпечує максимальну ефективність.

Іншим критичним фактором у центрах обробки даних є те, що витрати на технічне обслуговування апаратного забезпечення завжди дотримуються при використанні систем ШІ для керування всім центром. Центри обробки даних завжди вимагають захисту від факторів навколишнього середовища, оскільки вони містять важливу інформацію для клієнтів або організацій. Виходячи з цього фактора, штучному інтелекту завжди необхідно забезпечувати безпеку машин. З роками все більше компаній і організацій включили системи ШІ у свої центри обробки даних для підвищення безпеки та ефективності. Цей фактор демонструє вплив ШІ на кібербезпеку. Хоча штучний інтелект є корисним для кібербезпеки, існують інші обмеження, пов’язані з використанням штучного інтелекту в кібербезпеці.

Здатність змінювати те, що пропонує природа для сприяння діяльності та виживанню, завжди була метою людства, щоб забезпечити краще середовище для життя. Переходячи до індустріальної стадії людської революції, люди зробили внесок у забезпечення того, щоб широко використовувати знання про машини, які допоможуть у повсякденній діяльності [45]. Ідея знання фізики та способів використання та вдосконалення техніки допомогла людству повністю замінити тварин, дозволивши їм займатися своєю діяльністю. За допомогою обладнання вони змогли переконатися, що вони покращили свій продукт і ефективність у своїй роботі. Людина приходить, щоб зрозуміти, що техніка краща за людей. Таким чином, мета полягала в тому, щоб повністю перепланувати виробництво за допомогою машини, щоб отримати більш якісне виробництво та уникнути будь-яких незручностей, спричинених діями людини. І розробивши техніку, вони могли б отримати комп'ютерні технології, які ми маємо сьогодні.

Комп'ютерні технології стали однією з найбільш широко використовуваних технологій сьогодні, в результаті чого багато важливих елементів життя підтримуються технологіями. Таким чином, деякі стандарти повинні бути реалізовані в технології, щоб гарантувати, що ефективність і безпека пропонованих послуг викликають занепокоєння. Технологія належить фінансовим установам та іншим секторам, які зберігають важливу інформацію про наше життя. Крім того, технологія містить інформацію про нашу організацію, яку інші організації можуть використовувати для створення конкурентної переваги. Враховуючи, наскільки важливою є інформація для сучасного світу, комп'ютерні техніки та розробники повинні переконатися, що вони включили всі протоколи безпеки для забезпечення безпеки даних, задіяних у системі. Комп'ютерники повинні були розробити спосіб забезпечення безпеки даних; отже, вони повинні були зашифрувати свої дані перед їх надсиланням. Протокол шифрування гарантує, що якщо дані потраплять не тим людям, вони все одно не зможуть ними скористатися. Необхідно мати код дешифрування для декодування залучених даних, що

ускладнює його використання. Генерація шифрування даних тривала, щоб люди зрозуміли принципи, які використовуються в цьому процесі. На малюнку 5 нижче показано, як шифрування даних і бар'єри бізнес-процесів є перепоною для використання ШІ в усіх організаційних викликах, включаючи кіберзагрози для створення цінності.

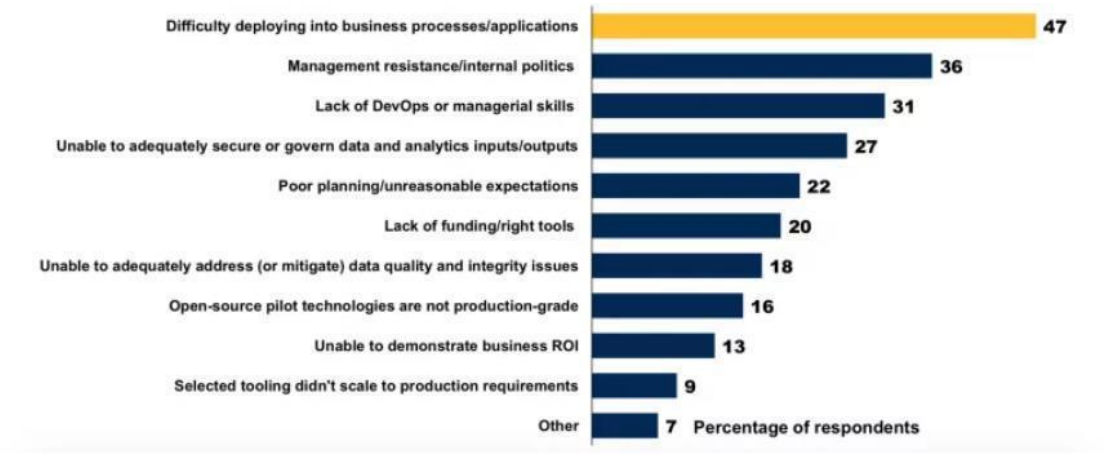

Рис. 3 Перешкоди на шляху впровадження ШІ проти кіберзагроз для досягнення цінності бізнесу

Люди, які навчаються та розуміють процес, дійшли до вирішення ще однієї проблеми. Люди, які вивчали протоколи, які використовуються системами та програмами шифрування, полегшили зворотне проектування процесу. Можливість отримати дані, що передаються, за допомогою протоколу ідентифікації ключа шифрування значно ускладнює весь процес захисту даних. Комп'ютерникам довелося розробити складніші протоколи та методи для шифрування даних і забезпечення їх правильної роботи. Кінцева мета захисту даних досягнута. З огляду на те, що машини кращі за людей у всьому, на що вони запрограмовані, логічно сказати, що вони найкраще забезпечуватимуть їх безпеку. Результатом цього є запровадження технології штучного інтелекту, яка забезпечує чудову безпеку машини. Система штучного інтелекту працює над тим, щоб вони призначили всі протоколи, на які вони запрограмовані, щоб гарантувати безпеку залучених даних.

ШІ використовує різні протоколи шифрування даних і використовує інші методи. Він може генерувати більш складний спосіб вирішення або

шифрування даних. За допомогою цих різних протоколів шифрування даних система може гарантувати, що буде достатньо важко гарантувати, що ніхто не зможе декодувати дані, задіяні в транзакції. ШІ ефективно обслуговує мережеві компанії та інші організації, оскільки безпека даних є більш просунутою та гарантованою. Однак, враховуючи, що люди створили технологію, у них є недоліки, навіть якщо вони були розроблені, щоб перепрограмувати та розвивати себе у разі будь-якої відповідальності. Той факт, що людина створила програму, дає їй можливість вивчити її та провести реверсивну інженерію залученого процесу, таким чином піддаючи проблему безпеки ризику потрапити в чужі руки.

Одним із найважливіших обмежень штучного інтелекту є те, що це лише комп'ютерний код, запрограмований для того, щоб переконатися, що вони дотримуються протоколів і розвиваються. Цей приклад може звучати добре, оскільки вони можуть розвиватися в будь-якому випадку. Однак система повністю запрограмована; тому будь-хто може взяти їх під контроль, ними можна маніпулювати та використовувати їх як зброю. Потрібно відредагувати кілька рядків коду, і тоді довгі робочі години можуть перетворитися на зброю, яку використають проти нього самого. Тому, маючи відповідні здібності та знання, технологію штучного інтелекту можна використовувати як зброю, яка використовуватиметься для знищення захисту, для якого вона створена. Цей фактор є одним із найсуттєвіших обмежень ШІ для кібербезпеки. Розробники та комп'ютерники повинні враховувати це, оскільки вони розуміють можливості технології ШІ.

Системи штучного інтелекту також можна навчити виявляти кіберзагрози та зловмисне програмне забезпечення, що робить їх ефективнішими в кібербезпеці. Зростаюча кількість атак на кібербезпеку призвела до впровадження ШІ в кібербезпеку. Весь процес полягає в тому, щоб забезпечити ефективність і точність. Однак штучний інтелект обмежений і не може замінити людей, оскільки йому доручено виконувати лише конкретне завдання. Часом він не може виявити практично нерозрізнені загрози і, отже, потрапляє в

проблеми, оскільки виглядає як справжнє повідомлення. ШІ також може бути важко виявляти загрози через розвиток кіберзагроз. Віруси та зловмисне програмне забезпечення вдосконалюються в будь-який момент часу, тому система штучного інтелекту має потребувати вдосконалення для підвищення ефективності. Крім того, практику кібербезпеки можна порівняти з кіберзлочинцями, які, як правило, отримують більше інформації про хакерство. Таким чином, кіберзлочинці можуть створити кращу загрозу, яку штучний інтелект не зможе легко виявити. Хоча штучний інтелект економить час для команди безпеки, він також вимагає людей-експертів для творчості, що полегшує їм роботу. Обмеження вимагає від розробників переконатися, що вони оснастили технологію кількома можливостями для боротьби з будь-якими злочинами, спричиненими їхніми обмеженнями.

З іншого боку, ми можемо визначити, що технологія ШІ не повністю використовується для захисту даних і забезпечення безпеки даних. Ми також можемо мати технологію штучного інтелекту, яка розроблена для генерування та створення комп'ютерних вірусів. Завдяки складності технології окремим особам важко конкурувати з машиною, що призводить до з'єднання даних. Оскільки коди, згенеровані штучним інтелектом, які використовуються для розробки комп'ютерного вірусу, створюють стільки можливостей, то цілком можливо пошкодити базу даних і маніпулювати даними в ній. Це ще одне обмеження, яке по суті стосується не технології ШІ як сторожового пса кіберзлочинності, а як учасника кіберзлочинності. Це обмеження демонструє ще один величезний вплив ШІ на кібербезпеку. Складність технології є обмеженням технології ШІ, оскільки не всі в суспільстві знають технологію. Існує також той факт, що зрозуміти різні моделі, задіяні в технології, непросто. Технологія, яка є настільки складною для використання та впровадження на повну потужність, може дати злочинцям шанс отримати систему, оскільки ми не можемо використовувати підхід до максимальної можливості. Технологія вимагає багато інформації про її роботу, якої багато людей не мають. Тому

організації все ще знаходяться в зоні ризику, оскільки вони не можуть забезпечити найкращу роботу системи.

Крім того, оскільки у нас є складність системи, ми розуміємо, що технологія коштуватиме багато. Тому вартість впровадження технології значно дорожча. Тому не всі організації у світі зможуть отримати доступ до технології та забезпечити безпеку даних. Тому вартість технології також є обмеженням для впровадження. Незважаючи на те, що інформація організації є важливою для будь-якої організації, вартість впровадження технології штучного інтелекту набагато вища, що обмежує кількість осіб, які використовуватимуть цю технологію для безпеки своїх даних та інформації. Вартість системи призводить до того, що небагато членів і організацій використовують цю технологію, тому важко оцінити можливості технології.

Висновки до розділу 2

В другому розділі було представлено використання ШІ для прогнозування та управління кіберризиками, таксономію зон фокусування (AoF) для механізму когнітивного зворотного зв'язку в прогнозній аналітиці кіберризиків та переваги і обмеження використання ШІ в кібербезпеці. Завдяки своїм можливостям аналізувати великі обсяги даних, виявляти складні закономірності та приймати автономні рішення, ШІ має потенціал значно підвищити ефективність кіберзахисту. Це звільняє час аналітиків кібербезпеки для зосередження на більш складних завданнях, таких як розслідування інцидентів та розробка стратегій захисту. ШІ може використовуватися для аналізу історичних даних про кібератаки та виявлення факторів, що сприяють їх виникненню. Ці знання можна використовувати для прогнозування майбутніх загроз та розробки превентивних заходів. Впровадження зон фокусування (AoF) підкреслює важливість конфіденційності в механізмах зворотного зв'язку, що використовуються для звітів про кібератаки та спільних баз даних в аналізі ризиків кіберфізичних систем (CPS). Переваги використання Штучного

Інтелекту в кібербезпеці в тому, що це покращює точність та швидкість виявлення кіберзагроз,а також автоматизація рутинних завдань, прогнозування та управління кіберризиками, та зниження витрат на кібербезпеку. А обмеження використання ШІ полягає в упередженості та дискримінації в алгоритмах ШІ що веде до потенційної вразливості до кібератак. Висока вартість розробки та впровадження систем ШІ, етичні проблеми, пов'язані зі збором та аналізом даних також не є винятком. Загальний вплив штучного інтелекту на кібербезпеку є як позитивним, так і негативним. З одного боку, ШІ пропонує потужні інструменти для боротьби з кіберзагрозами та покращення кіберзахисту. З іншого боку, важливо усвідомлювати потенційні обмеження та ризики, пов'язані з використанням ШІ в сфері кібербезпеки. Важливо зазначити, що ШІ не є панацеєю від усіх кіберзагроз. Ефективне використання ШІ в кібербезпеці потребує ретельного планування, впровадження та етичного підходу. Необхідно також поєднувати ШІ з традиційними методами кібербезпеки для забезпечення всебічного захисту.

Розділ 3 РИЗИКИ ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ В КІБЕРБЕЗПЕЦІ

3.1 Потенційні загрози безпеці внаслідок недоліків у розробці та використанні ШІ

Останніми роками, особливо у 2023 році, різні організації розширили виробництво методологій і посібників, щоб зосередити увагу на ризиках безпеки штучного інтелекту та допомогти компаніям успішно їх запобігати.

Таким чином, Агентство Європейського Союзу з кібербезпеки (ENISA) опублікувало різні рамки, методології та звіти, що стосуються ризиків безпеки штучного інтелекту та проблем, з якими доводиться стикатися компаніям.

Національний інститут стандартів і технологій США (NIST) створив структуру для управління ризиками безпеки ШІ. Фонд OWASP , глобальний методологічний еталон, запустив проект з усунення ризиків безпеки ШІ.

Що стосується регулювання, Європейський Союз робить останні кроки в обробці та затвердженні Закону про штучний інтелект. У проекті регламенту, який зараз мають обговорити Європейський парламент і Рада, наголошується на зв’язку між кібербезпекою та ШІ.

Нижче розглянуто основні ризики безпеки ШІ , які компанії, що розробляють штучний інтелект і використовують цю проривну технологію, повинні враховувати, щоб виявляти загрози, запобігати інцидентам безпеки та відповідати нормативній базі, яка ставатиме дедалі вимогливішою.

2023 рік залишиться в нашій пам’яті роком, коли генеративний ШІ, тобто той, хто здатний створювати контент і відповідати на запити людей, привернув увагу світової громадськості. ChatGPT, Midjourney, DALL-E, Copy AI та ін.

Однак історію штучного інтелекту можна простежити від Алана Тьюрінга, за десятиліттями досліджень і розробок машинного навчання, нейронних мереж, глибокого навчання та інструментів обробки природної мови.

Штучний інтелект уже присутній у багатьох пристроях і технологіях, які компанії та громадяни використовують щодня для автоматизації завдань або оптимізації прийняття рішень. У цьому сенсі штучний інтелект став чудовим союзником професіоналів у сфері кібербезпеки та компаній у зміцненні їхніх захисних можливостей і підвищенні стійкості до кібератак.

Не заходячи далі, завдяки інструментам машинного навчання стало можливим автоматизувати виявлення загроз або впровадження механізмів реагування на атаки. А також оптимізація оцінки безпеки та встановлення пріоритетів вразливостей, прогнозування моделей атак ворожих учасників, вилучення інформації з високою доданою цінністю з даних для виявлення вразливостей раніше, ніж це зроблять злочинці, або вдосконалення криміналістичного аналізу для усунення виявлених проблем.

Зі зростанням актуальності систем штучного інтелекту та їх потенціалу для компаній і громадян ризики безпеки ШІ стали серйозною проблемою з точки зору кібербезпеки .

Подібно до того, як штучний інтелект має вирішальне значення для розробки та реалізації послуг кібербезпеки, він має вирішальне значення, коли мова йде про захист систем і програм ШІ в той час, коли:

- Кіберзлочинці починають націлюватися на ШІ.
- Ризики, пов’язані з ланцюгом постачання, стають дедалі вищими.

Якщо є один фундаментальний елемент, коли мова йде про ШІ, особливо про системи машинного та глибокого навчання, то це дані. Ці системи працюють завдяки моделям, які потрібно навчати за допомогою даних.

Якщо дані численні, різноманітні, неупереджені та не маніпулювалися, модель може функціонувати оптимально та працювати на високому рівні. З

іншого боку, якщо дані, використані для навчання моделей, були пошкоджені, модель демонструватиме маніпуляцію, що може мати серйозні наслідки для компаній і користувачів.

Виходячи з цього, ми повинні зазначити, що, розглядаючи ризики безпеки ШІ, ми повинні розрізняти:

- Загрози, націлені на системи ШІ. Атаки, метою яких є системи ШІ: моделі, дані тощо.
- Зловмисне використання інструментів штучного інтелекту для здійснення кібератак на корпоративне програмне забезпечення, системи або окремих осіб.

Проект, запущений OWASP для оцінки головних ризиків безпеки та конфіденційності, з якими стикається штучний інтелект, визначає кілька небезпек і приділяє особливу увагу потенційним атакам на моделі ШІ.

Конвеєр ШІ слід розглядати як нову поверхню для атаки, оскільки він лежить за межами традиційної сфери розробки програмного забезпечення. чому Він включає науку про дані.

І інженерія даних, і інженерія моделей є важливими для розробки систем ШІ. Обидві дисципліни вимагають надійного контролю безпеки, щоб запобігти витоку чи отруєнню даних , крадіжці інтелектуальної власності чи атакам на ланцюг поставок.

З іншого боку, ризик пов’язаний з використанням даних під час розробки ШІ. Щоб навчити та протестувати модель, спеціалісти з обробки даних мають працювати з точними даними, які можуть стати конфіденційними. Отже, має бути реалізований суворий механізм контролю доступу, за якого спеціалісти з обробки даних можуть отримати доступ лише до інформації, необхідної для виконання своєї роботи.

Як зазначалося вище, атаки на модель штучного інтелекту життєво важливі для вирішення проблеми кібербезпеки штучного інтелекту. Ці високоризикові атаки можна запобігти:

- Захист процесу розробки ШІ.

- Приховування параметрів моделі.
- Обмеження доступу до моделі.
- Впровадження системи моніторингу для виявлення шкідливих даних.
- Розгляд такого роду атак під час фази навчання моделі.

Таким чином, необхідно поєднувати знання з кібербезпеки з навчанням машинного навчання. Крім того, можна запровадити класичні заходи кібербезпеки, такі як застосування принципу найменших привілеїв.

OWASP компілює наступні типи атак на моделі ШІ:

- Отруєння даними. Якщо навчальні дані змінюються, поведінкою моделі можна маніпулювати. Ця зміна дає змогу саботувати систему штучного інтелекту або змусити її приймати рішення, яких хоче зловмисник.
- Маніпуляція введенням. Ця атака спрямована на маніпулювання моделями з оманливими вхідними даними. Швидка ін'єкція є парадигмальним прикладом цього типу атаки.
- Висновок про членство. Запис даних і доступ до моделі за допомогою чорної скриньки можуть визначити, чи був документ у навчальному наборі даних. Це питання передбачає, що ворожі актори можуть знати, чи страждає особа на конкретну хворобу, чи є членом політичної партії чи певної організації.
- Інверсія моделі або реконструкція даних. Взаємодіючи з моделлю, оцінюються дані її навчання. Якщо ці дані конфіденційні, конфіденційність може бути порушена.
- Крадіжка моделі. Взаємодія з моделлю може призвести до визначення її поведінки та копіювання для навчання іншої моделі, що є крадіжкою інтелектуальної власності.
- Модель атаки на ланцюг поставок. Ці атаки можуть маніпулювати життєвим циклом моделі, наприклад, заражаючи базову модель, яка була оприлюднена, і пошкоджуючи моделі глибокого навчання, які використовують навчання передачі для вдосконалення цієї моделі.

Фахівці з обробки даних зосереджені на створенні робочих моделей і менше на створенні коду, який легко читати іншим фахівцям. Це рішення ускладнює аналіз коду ШІ, виявлення помилок або керування вразливими місцями. Отже, важливо поєднувати знання спеціалістів із обробки даних із підготовкою та досвідом інженерів програмного забезпечення та експертів з кібербезпеки.

Штучний інтелект робить ланцюжок постачання програмного забезпечення більш складним. По-перше, оскільки системи штучного інтелекту зазвичай мають кілька ланцюжків постачання (дані, моделі), джерела походження можуть бути паралельними або послідовними. Якщо до цього додати релевантність атак на моделі та їхню поведінку, яка не може бути оцінена за допомогою статичного аналізу, ми зіткнемося з дуже серйозним ризиком.

Таким чином, традиційну програмну специфікацію матеріалів (SBOM) необхідно доповнити специфікацією AI (AIBOM), одночасно вживаючи необхідних заходів для аудиту безпеки постачальників. Таким чином, управління ланцюгом поставок ШІ стає важливим аспектом безпеки ШІ.

Як і у випадку з традиційною розробкою програмного забезпечення, дослідники даних можуть отримати вигоду від відкритого вихідного коду, хоча він може містити слабкі місця та вразливості, які впливають на безпеку та конфіденційність. Тому необхідно проводити ретельний контроль повторно використовуваного коду.

На додаток до всіх атак, спрямованих на системи штучного інтелекту, цю проривну технологію можуть використовувати злочинці для оптимізації своїх можливостей атак. Іншими словами, ризики безпеки ШІ включають не лише загрози для систем . Вони також включають ШІ як інструменти на службі кіберзлочинців.

У своєму звіті про штучний інтелект і кібербезпеку ENISA наводить приклад складних кібератак, використовуючи шкідливий генеративний

штучний інтелект для створення глибоких фейків і маніпулювання інформацією, голосами, зображеннями, відео та навіть обличчями.

Але ми також повинні враховувати атаки, які вимагають менше ресурсів і знань. Наприклад, використання генеративного штучного інтелекту для створення переконливих текстів для атак на людей, компанії та установи за допомогою методів соціальної інженерії: фішинг, смішинг… Або використання штучного інтелекту для визначення того, які вразливості легше використати для атаки на корпоративні системи організації.

Подібним чином системи штучного інтелекту можуть оптимізувати ефективність і дієвість шкідливого програмного забезпечення, яке використовується групами кіберзлочинців, у кількох ключових аспектах: уникнення механізмів виявлення, адаптація до мінливого середовища, розповсюдження та стійкість у атакуваних системах.

Крім того, зловмисне програмне забезпечення на основі штучного інтелекту може використовувати методи навчання та підвищувати їх ефективність, здійснюючи більш успішні атаки.

Так само, як системи штучного інтелекту знаходяться в стадії повного розвитку, кібератаки, розроблені на основі потенціалу цієї технології, також швидко розвиваються. Таким чином, вони стануть більш досконалими в найближчі роки, а їхній потенційний вплив на компанії та громадян буде більш значним. Тому необхідно посилити стратегії кібербезпеки, враховуючи цей новий спектр атак.

ENISA розділила зловмисників на сім типологій із різними характеристиками та цілями:

- Кіберзлочинці. Групи кіберзлочинців мають чітку мету: отримання прибутку. Щоб отримати економічну вигоду від своєї злочинної діяльності, вони можуть використовувати системи ШІ для здійснення атак або атакувати ці системи напряму. Наприклад, злам чат-ботів штучного інтелекту для доступу до конфіденційної інформації, такої як банківські реквізити клієнтів компанії.

- Сценарій діточок. Цьому класу кіберзлочинців бракує знань для здійснення складних атак і розробки шкідливого програмного забезпечення, тому вони використовують упаковані інструменти атак і заздалегідь написані сценарії для атаки на корпоративні системи.
- Урядові актори та спонсоровані державою групи. Подумайте, наприклад, про групи APT, спонсоровані країною. Ці групи мають багато ресурсів і великий досвід, що дозволяє їм розробляти більш витончені та складні атаки. Їхні цілі можуть варіюватися від нападу на критично важливі сектори та інфраструктури країни до дестабілізації її демократичної системи шляхом зміни виборів і поширення дезінформації до викрадення конфіденційної інформації в компаній і державних адміністрацій.
- Терористи. Кібертерористи прагнуть завдати прямої шкоди життю людей, навіть спричинити смерть, наприклад, шляхом саботування ключових інфраструктур або чутливих секторів, таких як охорона здоров’я. Тероризм був постійним лихом протягом 21-го століття, і тепер він є не лише проблемою безпеки, але й проблемою кібербезпеки.
- Співробітники компанії та постачальники. Люди, які мають доступ до критично важливих елементів ШІ, таких як моделі або набори даних, можуть навмисно саботувати системи, наприклад, отруюючи навчальні дані. Крім того, вони також можуть ненавмисно викликати інциденти безпеки, випадково пошкодивши дані.
- Конкуруючі компанії . Конкуренція в технологічному секторі щодо штучного інтелекту зростає, тому атаки можуть надходити від конкуруючих компаній, які прагнуть викрасти інтелектуальну власність або підірвати репутацію компаній, що розробляють або використовують системи ШІ.

Хактивіст - ця концепція змішує хакерство з активізмом для позначення ворожих акторів, чия мотивація є переважно ідеологічною та які прагнуть зламати системи штучного інтелекту, щоб висвітлити їх вразливі місця та ризики.

Розвиток штучного інтелекту та популярність генеративного ШІ останнім часом посилили появу голосів, які попереджають про небезпеку ШІ не лише з точки зору кібербезпеки.

Чим ризики ШІ відрізняються від ризиків традиційного програмного забезпечення?

Системи штучного інтелекту є програмним забезпеченням, але з певними особливостями, які роблять їх більш складними та розширюють зону атаки традиційного програмного забезпечення.

Ось чому система управління ризиками ШІ, розроблена NIST, об'єднує деякі нові ризики, пов'язані з розвитком ШІ. NIST також детально описує інші загрози щодо програмного забезпечення, яке ми використовуємо щодня, але вони посилилися.

Деякі з цих ризиків не пов'язані безпосередньо з кібербезпекою. Наприклад, обчислювальні витрати на розробку ШІ, складність завдань обслуговування або вплив цих технологій на навколишнє середовище.

Однак ця структура розглядає ризики, пов'язані з безпекою систем, організацій і користувачів.

Нові та складніші виклики кібербезпеці

1. Дані, які використовуються для побудови системи штучного інтелекту, можуть невірно відображати контекст або передбачуване використання системи, а якість даних може вплинути на надійність штучного інтелекту з негативними наслідками.

2. Залежність систем ШІ від даних, які використовуються для навчання.

3. Зміни на етапі навчання, навмисні чи ненавмисні, можуть змінити продуктивність системи ШІ.

4. Набори даних, які використовуються під час навчання ШІ, можуть застаріти після розгортання системи, що вплине на продуктивність ШІ.

5. Існуюча невідповідність між масштабом і складністю систем штучного інтелекту та звичайних програмних додатків, які їх розміщують.

6. Попередньо підготовлені моделі мають вирішальне значення для сприяння дослідженню штучного інтелекту та розробці високопродуктивних систем за менший час і з меншими витратами. Однак вони також можуть збільшити рівень статистичної невизначеності та спричинити зміщення та проблеми з відтворюваністю .

7. Численні ризики для конфіденційності як наслідок величезної здатності систем штучного інтелекту агрегувати дані.

8. Виконувати тестування безпеки програмного забезпечення на основі штучного інтелекту складніше, оскільки розробка коду штучного інтелекту відрізняється від традиційної розробки коду, і можуть виникнути запитання щодо того, що і як тестувати.

Ризики безпеки, характерні для систем штучного інтелекту, вимагають від організацій запровадити стратегії для управління ризиками кібербезпеки та конфіденційності на всіх етапах життєвого циклу штучного інтелекту: проектування, розробки, розгортання, оцінки та використання.

У зв’язку з цим послуги кібербезпеки мають бути включені до програм безпеки компаній. Моделювання загроз, аналіз ризиків, навчання професіоналів, статичний і динамічний аналіз, аналіз коду, пентестування та вправи Red Team.

Ці завдання посилять безпеку, стійкість і конфіденційність систем ШІ, щоб досягти наступного:

- Захист додатків та IT-інфраструктури , приховування параметрів моделі для захисту від атак.
- Посилити захист нових конвеєрів розробки, пов’язаних із ШІ.
- Належним чином керуйте проблемами упередженості в системах ШІ.
- Звернути увагу на ризики, пов’язані з генеративним ШІ.
- Вирішуйте проблеми безпеки, пов’язані з ухиленням, вилученням моделі, доступністю системи та впровадженням атак машинного навчання.

- Аналізуйте та відстежуйте складну поверхню атак систем штучного інтелекту, щоб виявляти атаки на ранніх стадіях Cyber Kill Chain і дивну поведінку.
- Розгляньте ризики, пов’язані з технологіями ШІ, розробленими третіми сторонами.

Дослідження штучного інтелекту в розпалі. Технологічні інновації в цій галузі, а також розробка програмного забезпечення повинні сприяти підвищенню надійності та продуктивності систем, а також їхньої безпеки та стійкості до дій ворожих гравців.

Безпека ШІ є важливою проблемою в цю епоху

Впровадження штучного інтелекту в різні виробничі сектори та демократизація доступу до ШІ з інструментами, доступними для малих і середніх підприємств, а не лише для великих компаній, є новою віхою в технологічній революції, яку ми пережили в останні десятиліття.

Таким чином, ризики безпеки штучного інтелекту повинні бути центральними для публічних дебатів і серцевиною бізнес-стратегій.

Успішні атаки на системи ШІ можуть мати катастрофічні наслідки для компаній, які їх розробляють, а також для компаній, які їх використовують, і громадськості: викрадання приватних даних, дезінформація, втрата репутації, правові наслідки.

Тому важливо займатися безпекою штучного інтелекту протягом усього його життєвого циклу, запроваджуючи адекватні засоби контролю безпеки та проводячи постійні оцінки безпеки.

Безпека за проектом і протягом усього життєвого циклу

Як зазначено в посібнику ENISA щодо ШІ та кібербезпеки, концепція безпеки за проектом , яка широко використовується в розробці програмного забезпечення, повинна бути перенесена на арену ШІ.

Шляхом інтеграції засобів керування кібербезпекою, механізмів і найкращих практик на ранніх етапах проектування та розробки систем

штучного інтелекту, додатків та IT-інфраструктур, які їх підтримують. Таким чином, агентство ЄС рекомендує:

1. Проводьте тестування безпеки та моделюйте загрози для виявлення вразливостей, недоліків і векторів атак.

2. Взяти на себе зобов'язання дотримуватися методів безпечного кодування та виконувати аудит вихідного коду для виявлення помилок і вразливостей.

3. Впроваджуйте безпечні методи обробки даних, щоб забезпечити конфіденційність і запобігти пошкодженню чи видобутку даних.

4. Виконуйте тести безпеки, щоб виявити проблеми безпеки на ранніх етапах процесу розробки. Такі тести, як DAST, необхідні для керування динамічними ризиками кібербезпеки та визначення пріоритетів загроз.

5. Переконайтеся, що системи штучного інтелекту розроблені прозоро, а їх поведінку можна постійно перевіряти, щоб виявити аномальну поведінку та виправити її, перш ніж вона призведе до інцидентів безпеки.

Коротше кажучи, ризики безпеки штучного інтелекту повинні розглядатися з максимальною ефективністю та суворістю, розумно застосовуючи всі багатства знань, передових практик, тестів і методологій, розроблених фахівцями з кібербезпеки за останні десятиліття.

З цією метою важливо запровадити засоби контролю безпеки з першої фази життєвого циклу системи штучного інтелекту, мати міждисциплінарні команди на місці та здійснювати комплексний моніторинг ланцюжка поставок штучного інтелекту. ШІ має важливе значення для оптимізації послуг кібербезпеки.

3.2 Виклики в області правового регулювання та етичної оцінки використання Штучного Інтелекту в кібербезпеці

Міжнародне співтовариство ще не виробило єдиного підходу до правового регулювання ШІ, і різні країни пропонують різні рішення.

У цьому підрозділі я розглянув різні підходи до правового регулювання штучного інтелекту в розвинених країнах і окреслив основні аспекти правових питань, які слід враховувати при інтеграції штучного інтелекту в бізнес-процеси. Важливо відзначити, що розробка ефективного та збалансованого правового регулювання ШІ залишається актуальним завданням для всього світового співтовариства.

Наразі не існує загальноприйнятого правового визначення поняття «штучний інтелект» серед держав і наукового співтовариства. Наприклад, у 2020 році в США було запропоновано таке визначення: штучний інтелект – це машинна система, здатна робити прогнози, рекомендації та рішення, що впливають на фізичне або віртуальне середовище, для певного набору завдань, визначених людиною.

Подібне визначення штучного інтелекту пропонується в законопроекті Європейського Союзу про штучний інтелект (Закон про штучний інтелект): AI — це програмне забезпечення, здатне генерувати такі результати, як контент, прогнози, рекомендації або рішення, що впливають на навколишнє середовище, у рамках завдань, визначених людиною.

Штучний інтелект насправді є системою, яка використовує комп’ютери та машини для збору й аналізу даних, щоб імітувати здатність людини робити прогнози, давати рекомендації та приймати рішення. На сьогоднішній день розробники технологій і дослідники, що працюють у сфері штучного інтелекту, ще не прийшли до консенсусу щодо того, якими навичками повинен володіти ШІ, щоб вважатися штучним інтелектом. Це пояснюється відсутністю точного визначення того, що таке мислення та інтелект у найзагальнішому розумінні.

Дослідники пропонують різні погляди на це питання. Деякі стверджують, що важливою характеристикою інтелекту є здатність створювати нову інформацію, нові ідеї та робити висновки, які не були однозначними під час навчання. Інші вказують на необхідність для ШІ мати внутрішню модель зовнішнього світу та здатність адаптуватися до навколишнього середовища.

Незважаючи на різноманіття думок, все більше дослідників погоджуються, що одним із ключових атрибутів ШІ є його здатність навчатися на основі інформації, отриманої з різних джерел, таких як Інтернет, зовнішні датчики та інші пристрої. Машинне навчання стало важливою частиною розвитку штучного інтелекту, що передбачає здатність системи навчатися на подібних завданнях.

Слід зазначити, що питання про те, чи є у штучного інтелекту свідомість, залишається предметом дискусій і досліджень. Деякі вчені припускають, що свідомість є кінцевою метою розробки штучного інтелекту, але на даний момент це питання залишається невирішеним.

Ровена Родрігес у своїй статті під назвою «Юридичні проблеми та проблеми з правами людини щодо штучного інтелекту: прогалини, виклики та вразливі місця» аналізує правові проблеми та виклики, пов’язані з впливом штучного інтелекту на основні права людини[47]. Він охоплює такі сфери:

- Відсутність прозорості в алгоритмах.
- Уразливості кібербезпеки.
- Питання несправедливості, упередженості та дискримінації.
- Відсутність можливості оскаржити результати алгоритму.
- Питання «особистості» (правосуб'єктності).
- Питання інтелектуальної власності.
- Питання негативного впливу ШІ на працю та зайнятість.
- Питання конфіденційності та захисту даних.
- Питання відповідальності за шкоду та збитки.

Штучний інтелект залежить від кількох факторів:

- Технічний фактор: якість алгоритмів, їх вразливість і захист.
- Соціальний фактор: рівень обізнаності населення про AI-технології, готовність до нових технологій, правові позиції.
- Політичний фактор: розвиток політичних концепцій щодо ШІ та управління ним.

- Регулятивний фактор: законодавство, моніторинг та юридична відповідальність.
- Економічний фактор: наявність ресурсів для усунення негативних наслідків, інвестиції в безпечні та етичні системи ШІ, рівень доходу та страхування.

Звертаю увагу на те, що сфера штучного інтелекту наразі не має достатнього рівня правового регулювання, тому виникає необхідність регуляторних органів, для більш чіткого застосування існуючого законодавства та більшого наукового розвитку в цій галузі. Також зазначу, що виклики, створені конвергенцією штучного інтелекту, робототехніки та технологій Інтернету речей, можуть бути ще складнішими та серйознішими в майбутньому.

Штучний інтелект – це галузь, яка зосереджена на створенні систем і машин, здатних виконувати завдання, які зазвичай потребують людського інтелекту. Проблеми штучного інтелекту включають розробку етичних систем, прозорість машинних рішень і розуміння того, як ШІ може вплинути на суспільство. З моменту появи штучного інтелекту та революційних досягнень, які він приніс, виникло хвилювання від людей, які бояться штучного інтелекту, до тих, хто непомітно впроваджує його в їх повсякденне життя. Завдяки своїй здатності навчатися на основі даних і приймати рішення ШІ створює низку етичних і практичних проблем. З одного боку, штучний інтелект обіцяє інноваційні рішення в таких сферах, як медицина, промисловість, освіта тощо. Це може підвищити ефективність, автоматизувати завдання та надати нові можливості. Однак його прийняття піднімає фундаментальні етичні питання, такі як конфіденційність даних, алгоритмічна дискримінація та упереджене прийняття рішень.

Серед галузей інформатики штучний інтелект є фундаментальною областю. Тим не менш, існують інші дуже актуальні та актуальні сфери, такі як кібербезпека. Кібербезпека стосується практики захисту комп’ютерних систем, мереж і даних від кіберзагроз, таких як хакерські атаки, зловмисне програмне

забезпечення та крадіжка інформації. Проблеми кібербезпеки включають збереження конфіденційності, цілісності та доступності інформації та захист конфіденційності користувачів. Зі зростаючою залежністю від технологій кібербезпека стала критичною проблемою в епоху цифрових технологій. У контексті кібербезпеки захист від кіберзагроз стає надзвичайно важливим. Взаємозв'язок пристроїв і онлайн-систем створює вразливі місця, якими можуть скористатися кіберзлочинці. Кібербезпека прагне гарантувати, що наші системи та дані захищені від атак, але також створює проблеми з точки зору захисту конфіденційності та забезпечення справедливості в доступі до технологій.

Розвиток штучного інтелекту дозволив створити методи пом'якшення ризиків безпеці в комп'ютерних системах. Однак, так само як штучний інтелект дозволив розробити нові методи для підвищення кібербезпеки, його неетичне використання також є потенційним інструментом для кіберзлочинців і кібератак. В цьому підрозділі розглянуто перетин етики, штучного інтелекту та кібербезпеки, аналізуючи етичні принципи, пов'язані зі штучним інтелектом, і проблеми безпеки, пов'язані з його використанням. Крім того, розглядається вплив цих питань на суспільство та те, як етика та етичне прийняття рішень відіграють фундаментальну роль у цьому контексті.

Кожного разу, коли виникає галузь науки, яка має такий значний вплив на життя людей, у роботі в цій галузі з'являються етичні принципи, яких необхідно поважати. Штучний інтелект не є винятком із цього; у сфері штучного інтелекту етика відіграє фундаментальну роль у формулюванні вказівок і принципів, якими керується розробка та впровадження систем з використанням методів ШІ.

Були запропоновані етичні принципи штучного інтелекту, і їх формулювання продовжують вдосконалюватися з роками, і ці принципи є важливими для забезпечення відповідального використання ШІ. Є багато етичних принципів, які були сформульовані в контексті ШІ; ми наводимо деякі дуже релевантні:

• Прозорість і зрозумілість: одним із фундаментальних принципів є прозорість у процесах прийняття рішень щодо ШІ. Системи штучного інтелекту повинні мати можливість пояснити, як вони приходять до своїх висновків, дозволяючи користувачам розуміти їхні рішення та довіряти їм. Це особливо важливо в критично важливих програмах, таких як охорона здоров'я та прийняття юридичних рішень.

• Справедливість і недискримінація: штучний інтелект не повинен підтримувати чи посилювати упередження чи дискримінацію, що існують у суспільстві. Системи штучного інтелекту слід навчати за допомогою різноманітних і репрезентативних даних, щоб гарантувати відсутність гендерних, расових або інших характерних упереджень у результатах. Необхідно вжити заходів для виправлення та запобігання дискримінації.

• Конфіденційність даних: захист конфіденційності даних є важливим. Системи ШІ повинні безпечно обробляти дані та поважати конфіденційність людей. Це передбачає дотримання таких нормативних актів, як Загальний регламент захисту даних (GDPR) у Європейському Союзі.

• Підзвітність і відповідальність: має бути встановлена відповідальність за рішення, прийняті системами ШІ. Це включає визначення того, хто несе відповідальність у разі негативних результатів або шкоди, заподіяної ШІ. Чіткість відповідальності має вирішальне значення для вирішення юридичних та етичних питань.

• Благодійність і нешкідливість: системи штучного інтелекту повинні прагнути приносити користь людству та уникати заподіяння шкоди. Це означає, що розробники та користувачі ШІ повинні ретельно розглянути етичні наслідки його використання та вжити заходів для мінімізації ризиків.

• Співпраця та людський нагляд: ШІ не повинен повністю замінювати людський нагляд і прийняття рішень. Натомість його слід використовувати як інструмент співпраці, який покращує процес прийняття рішень людьми, особливо в таких критичних сферах, як медицина та безпека.

- Безпечна розробка: кібербезпека є важливою у розробці систем ШІ. Системи повинні бути захищені від атак і вразливостей, щоб запобігти можливим негативним наслідкам.
- Справедливий розподіл переваг: необхідно забезпечити, щоб переваги штучного інтелекту були справедливо розподілені в суспільстві, а не зосереджені в кількох. Це передбачає розгляд таких аспектів, як справедливий доступ до технологій і навчання.
- Взяття на себе відповідальності за рішення, прийняті штучним інтелектом, є головним викликом у сфері безпеки та етики. Оскільки штучний інтелект стає більш автономним і приймає рішення, які впливають на різні аспекти суспільства, вкрай важливо визначити, хто несе відповідальність за відповідні дії та наслідки. Деякі ключові аспекти цього виклику:
- Відсутність нормативно-правової бази: оскільки штучний інтелект продовжує розвиватися, багато юрисдикцій ще не встановили чітких законів і правил щодо відповідальності за рішення щодо штучного інтелекту. Це може призвести до невизначеності та лазівок у разі аварій або проблемних ситуацій.
- Спільна відповідальність: у деяких випадках відповідальність за рішення ШІ може бути розподілена між кількома учасниками, включаючи розробників, власників систем, постачальників даних і користувачів. Визначення конкретних обов'язків кожної сторони може бути складним.
- Зміна в природі помилки: хоча людські помилки можуть бути більш зрозумілими та легшими, помилки штучного інтелекту можуть бути результатом складної та неінтуїтивної взаємодії між системою та навчальними даними. Це ускладнює визначення відповідальності в разі виникнення помилки.
- Рішення, пов’язані з життям і смертю: у критично важливих програмах, таких як автономні транспортні засоби чи системи охорони здоров’я, рішення, прийняті ШІ, можуть мати прямі наслідки для життя людей. Визначення відповідальних у таких ситуаціях є особливо складним і делікатним.

Щоб вирішити цю проблему, важливо, щоб розробники, користувачі та регулятори працювали разом, щоб створити чіткі правові та етичні рамки, які визначають відповідальність за використання ШІ. Крім того, важливо заохочувати прозорість у розробці алгоритмів і систем ШІ, щоб полегшити розуміння їхніх рішень і забезпечити вжиття заходів для пом'якшення упередженості та помилок. Етика та підзвітність у розробці та розгортанні ШІ мають вирішальне значення для використання його переваг і мінімізації пов'язаних ризиків.

Для забезпечення етики у сфері кібербезпеки фахівці з кібербезпеки часто дотримуються етичних кодексів. Етичні кодекси кібербезпеки — це документи, які встановлюють принципи та етичні стандарти, яких професіонали повинні дотримуватися під час виконання своєї професії. Ці кодекси спрямовані на сприяння етичній поведінці, відповідальності та доброчесності в поводженні з інформацією та безпекою систем. Незважаючи на те, що вміст може відрізнятися залежно від організації чи суб'єкта, який його випускає, можна включити деякі загальні моменти:

- Конфіденційність: фахівці з кібербезпеки повинні захищати конфіденційну інформацію, до якої вони мають доступ під час своєї роботи, і не розголошувати її без належного дозволу.
- Цілісність даних: вони повинні забезпечувати точність і точність даних, які вони обробляють, уникаючи несанкціонованих змін.
- Наявність систем: спеціалісти повинні переконатися, що системи та служби доступні та функціонують належним чином для авторизованих користувачів.
- Професійна відповідальність: вони повинні виконувати свою роботу сумлінно та компетентно, діючи в інтересах клієнта чи організації.
- Відповідність законам і нормативним актам: вони повинні поважати та дотримуватися законів і нормативних актів, що діють у їхній сфері діяльності.

• Конфлікт інтересів: уникайте ситуацій, які можуть спричинити конфлікт інтересів або поставити під загрозу їхню об’єктивність і неупередженість суджень.

• Не порушуйте конфіденційність: не вживайте дій, які посягають на конфіденційність людей, окрім випадків суворої необхідності та відповідно до чинного законодавства.

• Співпраця та відповідальне розкриття інформації: сприяйте відповідальному обміну інформацією про вразливості та загрози в спільноті кібербезпеки для покращення загальної безпеки.

• Повага до авторських прав та інтелектуальної власності: не використовуйте, не копіюйте та не поширюйте програмне забезпечення чи інші матеріали, захищені авторським правом, без належного дозволу.

• Прозорість: будьте прозорими та чесними у спілкуванні з клієнтами та роботодавцями щодо ризиків та обмежень систем безпеки.

Фахівці з кібербезпеки, як правило, зобов’язані переглядати та дотримуватися етичних кодексів, встановлених їх роботодавцем, професійною асоціацією чи відповідним регуляторним органом. Передача етичних принципів штучному інтелекту, який виконує завдання кібербезпеки, є важливим процесом, який гарантує, що ШІ діє відповідально, поважаючи конфіденційність і безпеку систем. Ось кілька способів донести ці етичні принципи до ШІ:

• Етичний дизайн: етичні принципи повинні бути включені з самого початку розробки ШІ. Інженери та розробники повинні враховувати етичні наслідки дизайнерських рішень, такі як конфіденційність, прозорість і повага до прав користувачів.

• Етичний набір даних: під час навчання ШІ слід використовувати набір даних, який відображає етичні принципи та не містить упередженості та дискримінації. Це не дасть ШІ навчитися небажаної або дискримінаційної поведінки.

• Рекомендації щодо поведінки: ШІ має бути запрограмований за допомогою алгоритмів і вказівок, які відображають етичні принципи. Наприклад, можуть існувати чіткі правила захисту конфіденційності даних і дотримання законів і правил.

• Моніторинг і аудит. Важливо відстежувати та перевіряти поведінку ШІ в режимі реального часу, щоб переконатися, що вона відповідає встановленим етичним принципам. Це передбачає моніторинг її діяльності та вжиття заходів у разі виникнення етичних проблем.

• Безперервне навчання та адаптація: штучний інтелект повинен бути готовий до навчання та адаптації на основі нових етичних принципів або змін у правилах і нормах.

• Взаємодія з людьми: якщо штучний інтелект взаємодіє з людьми, він повинен робити це етично та з повагою. Це може передбачати встановлення чітких обмежень щодо типу інформації, яку ШІ може збирати, і того, як він може її використовувати.

• Етичне тестування: перед впровадженням ШІ у виробниче середовище важливо піддати його етичному тестуванню, щоб оцінити його поведінку та виявити потенційні етичні проблеми.

• Прозорість: системи штучного інтелекту мають бути розроблені так, щоб їхні рішення та дії були зрозумілі людям. Прозорість забезпечує більшу підзвітність і полегшує виявлення упередженості чи невідповідної поведінки.

• Підзвітність розробників: розробники та менеджери штучного інтелекту повинні нести відповідальність за те, щоб ШІ діяв відповідно до встановлених етичних принципів.

• Навчання з питань етики. Команди розробників і фахівці з кібербезпеки, які працюють з ШІ, повинні пройти навчання з питань етики, щоб зрозуміти проблеми та етичні наслідки своїх дій.

Серед законів і нормативних актів, які існують і повинні дотримуватися як фахівцями з кібербезпеки, так і ІА, розробленими для цих завдань, є наступні:

- Закони про захист даних: у багатьох країнах діють закони про захист даних, які визначають, як слід обробляти та захищати персональні дані. ШІ, який використовується в кібербезпеці, повинен відповідати цим правилам, щоб забезпечити приватність і конфіденційність інформації.
- Закони про кібербезпеку. Деякі країни мають спеціальні закони, які регулюють безпеку комп'ютерних систем і мереж. Ці закони можуть вимагати певних стандартів безпеки та безпеки для організацій і компаній, що працюють у цій країні .
- Закони про відповідальність: штучний інтелект, який використовується в кібербезпеці, повинен підпадати під дію законів про відповідальність і цивільну відповідальність у разі порушення безпеки або інциденту, що впливає на треті сторони.
- Закони про інтелектуальну власність: Норми інтелектуальної власності повинні застосовуватися до алгоритмів і технологій ШІ, що використовуються в кібербезпеці, захищаючи права власників таких технологій.
- Положення для окремих галузей: деякі конкретні галузі, наприклад фінанси чи охорона здоров'я, можуть мати додаткові правила, які мають впливати на використання штучного інтелекту в кібербезпеці.

Описані закони не стосуються виключно експертів з кібербезпеки та систем ШІ, розроблених для цієї мети; швидше, вони охоплюють ту саму правову базу, що застосовується до кіберзлочинців. Ця правова парадигма поширюється на різні країни, де кримінальні кодекси окреслюють спектр кіберзлочинів, забезпечуючи однакове застосування принципів і норм, призначених для боротьби з незаконною діяльністю в цифровій сфері. По суті, ці закони служать комплексною основою, що виходить за межі відмінностей між тими, хто захищає цифрову безпеку, і тими, хто бере участь у кіберзлочинствах, встановлюючи стандартизований підхід до боротьби з кіберзагрозами в глобальному масштабі.

У середовищі кіберзлочинності, де таємні дії досвідчених хакерів часто вислизають від ідентифікації, Будапештська конвенція про кіберзлочинність постає ключовою відповіддю на ці виклики. Цей міжнародний договір, розроблений насамперед для європейських країн, окреслює всеосяжну структуру боротьби з кіберзагрозами та правопорушеннями. Він встановлює набір правових заходів і механізмів співпраці для боротьби з різними формами кіберзлочинності. Примітно, що його вплив виходить за межі Європи, оскільки дослідження та аналізи проводяться в таких країнах, як Перу та Чилі в Латинській Америці, що відображає глобальну актуальність його принципів. Серед злочинів, класифікованих у цих країнах, є такі:

- Несанкціонований доступ до комп'ютерних систем
- Комп'ютерне шахрайство
- Сприяння засобам комп'ютерного шахрайства
- Комп'ютерне шпигунство
- Несанкціоноване розголошення особистих даних або вмісту
- Отримання комп'ютерних даних
- Поширення програм, спрямованих на пошкодження або переривання
- Пошкодження комп'ютерної або телематичної систем
- Порушення, викрадення та видалення листування
- Незаконне перехоплення, перешкоджання або переривання комунікацій
- Фальсифікація, зміна або видалення вмісту комп'ютера
- Зберігання дитячої порнографії

Незважаючи на існування чітко визначених статутів і законів про кіберзлочинність, є випадки, коли досвідчений хакер уміло уникає залишати сліди, огортаючи свою особу завісою анонімності. Отже, встановлення особи, яка вчинила напад, або особи, яка порушила закон, стає надзвичайно складним завданням. У таких випадках відсутність явної підзвітності підкреслює невловиму природу цих досвідчених осіб, змушуючи експертів із

правоохоронних органів та кібербезпеки боротися зі складнощами атрибуції в цифровому ландшафті.

Важливо зазначити, що правове та нормативне середовище навколо штучного інтелекту постійно розвивається та може відрізнятися залежно від країни та регіону. Нові нормативні акти щодо штучного інтелекту в кібербезпеці могли з'явитися після того, як мені стало відомо. Тому я рекомендую вам звернутися до оновлених та юридичних джерел для отримання більш точної та актуальної інформації щодо чинних норм у цій сфері.

Кібербезпека з ШІ представляє низку етичних проблем і обмежень, які можуть вплинути на її ефективність і застосування. Деякі способи, якими етичні обмеження можуть впливати на методи кібербезпеки ШІ, включають наступне:

- Обмеження щодо збору та використання даних: етичні обмеження щодо збору та використання персональних даних можуть обмежити кількість і якість даних, доступних для навчання моделей ШІ. Це може вплинути на здатність систем ШІ ефективно виявляти та пом'якшувати загрози.
- Упередженість і дискримінація. Етичні міркування щодо уникнення упередженості та дискримінації в системах штучного інтелекту можуть призвести до обмеження або необхідності коригувати навчання з урахуванням історичних даних. Це може вплинути на точність і ефективність моделей у певних контекстах.
- Прозорість і пояснюваність: вимоги щодо прозорості та пояснюваності моделей штучного інтелекту можуть обмежити використання більш складних і важких для інтерпретації методів штучного інтелекту. Можна віддати перевагу простішим, але менш точним підходам для полегшення розуміння та підзвітності.
- Дилема безпеки проти конфіденційності: у деяких випадках може виникнути етична дилема між забезпеченням безпеки та захистом особистих даних. Необхідність зберігати конфіденційність і захищати права людей може

обмежити певні методи кібербезпеки, пов’язані з аналізом конфіденційних даних або доступом до них.

• Обмеження щодо експериментів і тестування. Етичні обмеження можуть ускладнювати тестування в реальному середовищі або експериментувати з потенційними загрозами для оцінки ефективності методів ШІ. Це може призвести до зниження рівня довіри до систем кібербезпеки штучного інтелекту до впровадження.

Таким чином, етичні обмеження можуть вплинути на розробку та впровадження інтелектуальних методів кібербезпеки штучного інтелекту, оскільки етичні цінності та принципи необхідно брати до уваги під час прийняття рішень щодо розробки, навчання та використання цих систем. Однак важливо також зазначити, що етичні міркування є важливими для того, щоб ШІ в кібербезпеці використовувався відповідально, чесно та з повагою до прав залучених користувачів і окремих осіб.

Завдяки дотриманню етичних обмежень і здорових етичних принципів зміцнюється довіра до програм ШІ в кібербезпеці. Довіра є вирішальним фактором у прийнятті та прийнятті цих технологій. Етичні обмеження встановлюють чітку відповідальність за розробку та використання ШІ в кібербезпеці. Це допомагає гарантувати, що організації та окремі особи несуть відповідальність за свої дії та рішення в разі виникнення інцидентів або проблем. Безвідповідальних або недбалих практик у впровадженні штучного інтелекту в кібербезпеку можна уникнути, а етичні проблеми можна виявити та вирішити до того, як вони стануть серйозними перешкодами. Це забезпечує більшу адаптивність і стійкість до систем кібербезпеки з підтримкою ШІ.

3.3 Стратегії мінімізації ризиків використання Штучного Інтелекту в Кібербезпеці

Щоб визначити можливі наслідки загроз, пов'язаних зі штучним інтелектом, і звести до мінімуму ймовірність їх реалізації, необхідно об'єднати

зусилля вчених, дослідників і розробників з боку вчених, науково-освітніх установ, промислових і промислових структур, державних і державних органів, законодавчої і виконавчої влади та міжнародного співтовариства.

Впровадження таких заходів допоможе створити систему зниження ризиків, яка дозволяє швидко виявляти, готуватися і реагувати на виклики і загрози, що виникають в результаті створення і використання систем штучного інтелекту.

Ризики, які створюють системи штучного інтелекту, є унікальними у багатьох відношеннях. Наприклад, системи штучного інтелекту навчаються даним, які можуть змінюватися з часом, а іноді і несподівано різко змінюватися, що може вплинути на функціональність і надійність системи способами, які важко зрозуміти. Системи штучного інтелекту та контекст, в якому вони розгортаються, часто складні, і їх важко виявити та реагувати на збої.

Системи штучного інтелекту носять соціотехнічний характер і залежать від соціальної динаміки і поведінки людини. Ризики та переваги штучного інтелекту можуть бути пов'язані з взаємодією технічних аспектів у поєднанні з соціальними факторами, пов'язаними з використанням системи.

1. Розвиток систем штучного інтелекту створює передумови для посилення існуючих загроз національній безпеці в інформаційній сфері. Ці загрози включають: Посилення кібератак. Штучний інтелект може підвищити ефективність процедур виявлення вразливостей в системах безпеки, виконання атак, маскування їх наслідків, імітацію поведінки людини в окремих фазах кібератаки.

2. Утворення каналів витоку інформації з обмеженим доступом. Системи штучного інтелекту можуть бути використані для посилення комп’ютерної розвідки завдяки аналізу шляхом аналізу великих обсягів даних, визначення трендів і патернів, щоб виявити конфіденційні дані про об’єкти, які стосуються національної безпеки, критичну інфраструктуру тощо. Зокрема, завдяки інтерактивній карті, що опублікована в Інтернеті та показує місцезнаходження

людей, які використовують такі фітнес-пристрої, як Fitbit, була продемонстрована можливість ідентифікації військових об’єктів США.

3. Атаки отруєння даних (*Data Poisoning,* DP) - це цільові атаки з метою модифікації або спотворення даних, що використовуються для машинного або глибокого навчання ШІ, внаслідок чого ШІ отримує небажані навички, які можуть завдати шкоди особі, суспільству та державі. Зазначимо, що процедури навчання ШІ зазвичай передбачають використання великої кількості даних для тренування моделі. Ці дані можуть бути зібрані з різних джерел і можуть містити помилки або неточності. Отруєнням даних - атака використовує ці неточності з метою введення помилкових чи зловмисних даних у навчальний набір.

Зважаючи на те, що ШІ можуть використовуватися для створення роботизованих збройних систем, які можуть самостійно визначати та атакувати цілі без участі оператора, DP атака може мати без перебільшення жахливі наслідки.

4. Містифікація даних. Спеціалістами компанії «Vulcan» виявлена схильність генеративної ШІ ChatGPT до створення недостовірних (галюцінованих) фактів і даних. А саме, ChatGPT в разі запиту рішень для кодування може пропонувати неіснуючі пакети (рис. 1, кроки 1, 2). Ці уявні пакети можуть бути використані зловмисниками для перетворення їх в замаскований шкідливий код, який завантажується в репозиторії кодів (крок 3). Якщо користувач, запитує у ChatGPT рекомендації щодо розробки (кроки 4, 5), в пропозиціях (кроки 6, 7) можуть з’явитися ці шкідливі пакети. Таким чином, ці недостовірні пакети СШІ потенційно перетворяться на канал витоку інформації користувач (крок 8).

Наведений вище перелік загроз не є вичерпним, але дає можливість оцінити складність, глибину та впливовість проблеми. Кожна з цих загроз вимагає глибокого розуміння технологій штучного інтелекту та формування ефективних стратегій для протидії їм.

Для блокування або нейтралізації визначених загроз необхідна реалізація низки заходів, спрямованих на мінімізацію (обробку) ризиків використання систем штучного інтелекту, і в першу чергу — на ідентифікацію ризиків використання ШІ. Як і ризики для інших типів технологій, ризики штучного інтелекту можуть виникати різними способами та можуть бути охарактеризовані як довготермінові чи короткострокові, з високою чи низькою ймовірністю, системні чи локалізовані, а також із сильним чи низьким впливом.

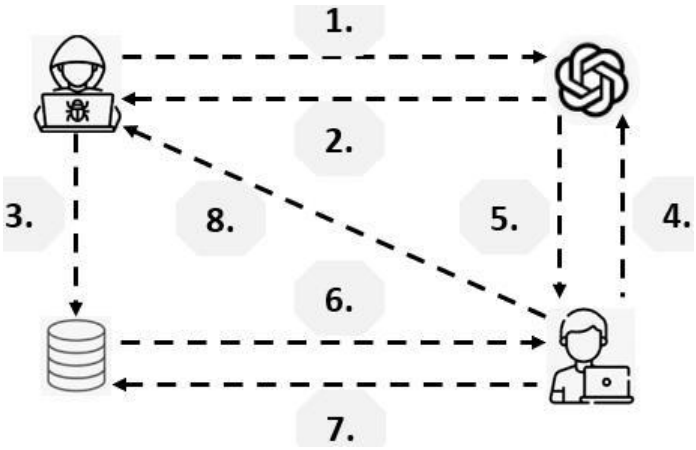

Рис. 4 Реалізація загрози витоку
інформації через недійсні пакети ШІ

Загальні принципи та орієнтири для управління ризиками будь-якого типу, розміру чи природи визначені стандартом ISO 31000:2018 Risk management — Guidelines [48].

Основна ідея полягає в тому, щоб створити систематичний і структурований підхід до ідентифікації, оцінки, керування та моніторингу ризиків. Ось деякі ключові етапи та аспекти управління ризиками за стандартом ISO 31000:

1. Встановлення контексту: Необхідно розуміти свій контекст і визначити фактори, що впливають на здатність до досягнення цілей;
2. Ідентифікація ризиків: Ризики ідентифікуються шляхом виявлення подій або ситуацій, які можуть впливати на досягнення цілей;

3. Оцінка ризиків: Визначення ймовірності та впливу ризиків для визначення їхньої значущості. Це допомагає визначити пріоритети для подальшого керування ризиками;
4. Обробка ризиків: В цьому етапі визначаються можливі стратегії обробки ризиків. Це може включати уникнення ризику, зменшення його впливу, передачу ризику і прийняття ризику;
5. Заходи управління ризиками: Розробляються та впроваджуються конкретні заходи для керування ризиками згідно з визначеними стратегіями;
6. Моніторинг та перегляд: Ризики та їхні заходи управління мають бути систематично переглянуті та оцінені для впевненості, що вони залишаються ефективними та актуальними;
7. Звітність та комунікація: Інформація про ризики та їхній стан повинна бути передана відповідним зацікавленим сторонам, включаючи керівництво та інші зацікавлені групи.

Загалом, управління ризиками за стандартом ISO 31000 вимагає системного підходу на всіх етапах життєвого циклу з області оцінювання.

Звернемо увагу, що в для нашого випадку фактично визначена мета управління ризиками: сучасні системи штучного зробити більш точними, безпечними, інтерпретованими, прозорими, надійними, узгодженими, заслуговуючими на довіру і лояльними.

Використання системи штучного інтелекту для управління ризиками, які виникають внаслідок використання інших ШІ може бути ефективним засобом забезпечення безпеки та стабільності. Такий підхід може включати в себе ряд етапів (рис. 5) та функціональних можливостей:

1. Ідентифікація ризиків: аналіз архітектури системи та алгоритмів, що використовуються, типи даних та інші параметри. Система може визначати слабкі місця, потенційні точки виникнення помилок або зони, де система може взаємодіяти з оточенням;

2. Моніторинг поведінки системи штучного інтелекту: система управління ризиками на базі системи штучного інтелекту може безперервно стежити за поведінкою інших систем управління, аналізуючи їх виходи, метрики та поведінку в реальному часі;
3. Автоматичне виявлення аномалій: використовуючи методи машинного навчання, система може виявляти аномалії або відхилення від норми в поведінці систем штучного інтелекту, що може свідчити про потенційні ризики;
4. Прогнозування ризиків: на основі історичних даних та актуального стану системи управління ризиками прогнозуються потенційні проблеми або непередбачувана поведінка системи штучного інтелекту у майбутньому;
5. Автоматична корекція: у випадках, коли виявлено ризик, система може автоматично вносити корективи в роботу іншої системи штучного інтелекту, наприклад, змінюючи її параметри або обмежуючи її дії;
6. Сценарії «чорної скриньки»: для вивчення і розуміння поведінки систем штучного інтелекту можна використовувати сценарії, де система штучного інтелекту піддається ряду тестів у контрольованому оточенні.
7. Аналіз причинно-наслідкових зв’язків: система може допомогти аналізувати причини певної поведінки системи штучного інтелекту, визначаючи, чи була ця поведінка результатом вхідних даних, алгоритмів, або інших факторів;
8. Зворотний зв’язок і навчання: на основі аналізу ризиків та інцидентів система може навчатися, вдосконалюючи свої методи виявлення та реагування на ризики.

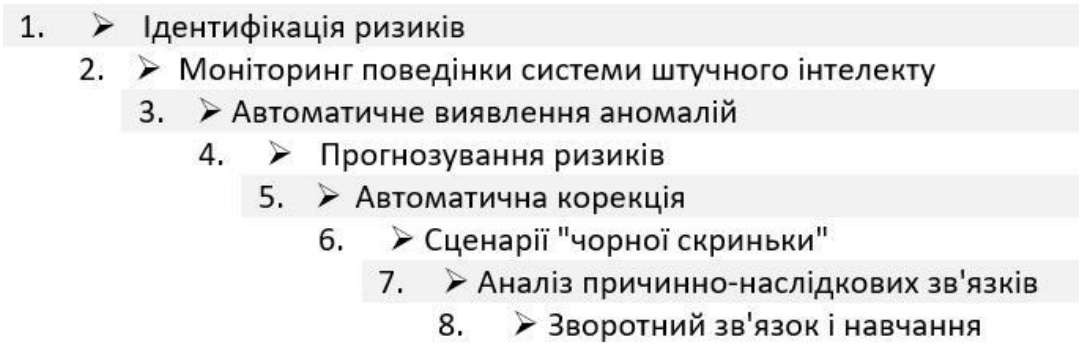

Рис. 5 Етапи використання ШІ для управління ризиками

Одним з найважливіших заходів є створення системи управління ризиками штучного інтелекту, на якому має базуватися регуляторна політика держави у цій галузі.

Система управління ризиками базується на використанні класичного підходу до оцінки ризиків, який наведено нижче:

1. Ідентифікація ризиків: Це початковий етап, де ідентифікуються потенційні ризики, пов’язані з ШІ. Це може включати розробку політик та алгоритмів, використання даних, вплив на користувачів і багато іншого;

2. Оцінка ризиків: Після ідентифікації ризиків вони повинні бути оцінені за їх потенційним впливом та ймовірністю виникнення. Це може включати аналіз чутливості, упередженості, моделювання ризиків або інші методики оцінки;

3. Прийняття рішень щодо ризиків: Після оцінки ризиків слід вирішити, як краще ними управляти. Це може включати прийняття рішень про вдосконалення процесів, модифікацію баз даних та баз знань ШІ або внесення змін в спосіб використання ШІ;

4. Управління ризиками: Це включає в себе виконання дій по управлінню ризиками, які було визначено на попередньому етапі. Це може включати виконання контрольних заходів, навчання персоналу, зміни в дизайні систем та інші заходи;

5. Моніторинг та перегляд ризиків: Ризики слід постійно моніторити та переглядати, щоб впевнитись, що вони залишаються під контролем та що вжиті

заходи ефективні. Це може включати регулярний аудит, моніторинг впливу, збір зворотного зв’язку від користувачів та інші механізми моніторингу.

Всі ці етапи мають повторюватись циклічно, оскільки ризики можуть змінюватися з часом. Також, оцінку ризиків необхідно проводити на всіх етапах життєвого циклу системи ШІ.

Реалізація заходів з управління ризиками ШІ пропонується за трьома напрямками: нормативно-правовим, технічним та організаційним.

Нормативно-правові заходи:

1. Визначення та прийняття державної політики в галузі штучного інтелекту;
2. Законодавче регулювання: Створення чітких законодавчих норм, які регулюють розробку та використання ШІ, може бути ефективним способом відповіді на ці загрози. Держава може прийняти закони, які обмежують використання ШІ в автономних збройних системах або встановлюють стандарти безпеки для ШІ в кібернетичних системах;
3. Міжнародне співробітництво: Участь в міжнародних угодах та ініціативах, спрямованих на регулювання ШІ, створенні міжнародних норм і стандартів безпеки для ШІ та забезпечення їх використання в Україні.

Технічні заходи:

1. Розробка безпечних систем ШІ: сприяння на рівні держави розробці та впровадженню безпечних систем ШІ, що включають вбудовані заходи безпеки;
2. Обмеження доступу до даних: Встановлення технічних обмежень на доступ ШІ до даних, таких як персональні дані громадян, що запобігає несанкціонованому використанню цих даних;
3. Створення систем ШІ для проведення аудиту знань прикладних систем ШІ;
4. Розробка механізмів виявлення ознак роботи небезпечних ШІ;
5. Розробка заходів з активної протидії небезпечним ШІ.

Організаційні заходи:

1. Управління ризиками: Створення моделі управління ризиками ШІ, яка містить механізми визначення рівнів загроз та імовірності їх реалізації в різних областях діяльності людини, суспільства, держави;

3. Освіта: Проведення освітніх кампаній для збільшення обізнаності про потенційні ризики, пов’язані з ШІ;

4. Співпраця з приватним сектором: Держава співпрацює з приватним сектором для створення безпечних систем ШІ і розробки ефективних стратегій протидії потенційним загрозам;

5. Створення спеціалізованих державних органів: Створення спеціалізованих органів, які будуть відповідальні за моніторинг та реагування на загрози, пов’язані з ШІ.

6. Організаційне обмеження доступу до масивів даних та спеціалізованих баз знань створених державними установами для використання їх в моделях навчання штучного інтелекту.

Найважливішим етапом в системі управління ризиками, які виникають в наслідок використання систем штучного інтелекту є оцінка ландшафту можливих ризиків та їх ідентифікація. Це ітеративний процес пошуку нових типів ризиків та профілювання їх основних характеристик для подальшої інтерпретації, аналізу та обробки. Завдання ідентифікації ризиків вирішується як завдання пошуку аномалій в масивах даних про діяльність, що стосується галузі застосування ризик-менеджменту. Аномальні спостереження в таких даних можуть пояснюватися наявністю взаємозв’язків та взаємодій між об’єктами та суб’єктами діяльності, що призводять до появи ще не ідентифікованих ризикових ситуацій та відповідних наслідків, або є потенційними джерелами виникнення таких ситуацій у майбутньому.

Висновки до розділу 3

В третьому розділі було досліджено потенційні загрози безпеці внаслідок недоліків у розробці та використанні ШІ, правове регулювання та етична оцінка ШІ, стратегії мінімізації ризиків використання ШІ в кіберпросторі. Однією з основних проблем є потенційні загрози безпеці через недоліки в розробці та застосуванні ШІ. Це може призвести до вразливостей, які можуть бути експлуатовані зловмисниками, включаючи автоматизовані атаки, зломи та витоки даних.

Виклики в області правового регулювання та етичної оцінки використання ШІ в кібербезпеці також є суттєвими. Правове регулювання має бути адаптоване до швидко змінюваних технологій ШІ, забезпечуючи адекватний захист прав та свобод громадян, одночасно сприяючи інноваціям. Етична оцінка використання ШІ потребує ретельного розгляду питань прозорості, відповідальності та довіри до систем ШІ. Необхідно забезпечити, щоб розробники та користувачі ШІ дотримувалися високих етичних стандартів, включаючи справедливість, неприйняття дискримінації та захист конфіденційності.

Стратегії мінімізації ризиків використання ШІ в кібербезпеці включають постійний моніторинг та оцінку алгоритмів, впровадження механізмів контролю та аудиту, а також навчання фахівців з кібербезпеки новим технологіям. Важливо розвивати міжнародне співробітництво для розробки спільних стандартів і практик використання ШІ, що сприятиме підвищенню загального рівня безпеки.

Отже, ефективне використання ШІ в кібербезпеці вимагає збалансованого підходу, що включає усунення технічних вразливостей, розробку належної правової бази та дотримання етичних принципів. Тільки таким чином можна забезпечити безпечне та надійне впровадження ШІ в кіберпростір.

ВИСНОВКИ

У даній дипломній роботі були розглянуті різні аспекти використання штучного інтелекту в кібербезпеці. Використання штучного інтелекту в кібербезпеці є однією з найбільш актуальних та обговорюваних тем сучасності. Ця робота розкриває перспективи і потенціал штучного інтелекту у галузі кібербезпеки. Ефективність використання штучного інтелекту в кібербезпеці полягає в його здатності виявляти, аналізувати та реагувати на кіберзагрози у реальному часі. Системи ШІ можуть автоматизувати процеси виявлення вразливостей, виявлення та блокування атак, що робить їх надзвичайно корисними для організацій у боротьбі зі зловмисниками в кіберпросторі.

Одним із найважливіших обмежень штучного інтелекту є те, що це лише комп’ютерний код, запрограмований для того, щоб переконатися, що вони дотримуються протоколів і розвиваються. Цей приклад може звучати добре, оскільки вони можуть розвиватися в будь-якому випадку. Однак система повністю запрограмована; тому будь-хто може взяти їх під контроль, ними можна маніпулювати та використовувати їх як зброю. А обмеження використання ШІ полягає в упередженості та дискримінації в алгоритмах ШІ що веде до потенційної вразливості до кібератак. Висока вартість розробки та впровадження систем ШІ, етичні проблеми, пов'язані зі збором та аналізом даних також не є винятком.

Наразі не існує загальноприйнятого правового визначення поняття «штучний інтелект» серед держав і наукового співтовариства. Сфера штучного інтелекту наразі не має достатнього рівня правового регулювання, тому виникає необхідність регуляторних органів, для більш чіткого застосування існуючого законодавства та більшого наукового розвитку в цій галузі

Завдяки дотриманню етичних обмежень і здорових етичних принципів зміцнюється довіра до програм ШІ в кібербезпеці. Довіра є вирішальним фактором у сприйнятті та прийнятті цих технологій. Етичні обмеження встановлюють чітку відповідальність за розробку та використання ШІ в

кібербезпеці. Це допомагає гарантувати, що організації та окремі особи несуть відповідальність за свої дії та рішення в разі виникнення інцидентів або проблем. Безвідповідальних або недбалих практик у впровадженні штучного інтелекту в кібербезпеку можна уникнути, а етичні проблеми можна виявити та вирішити до того, як вони стануть серйозними перешкодами. Це забезпечує більшу адаптивність і стійкість до систем кібербезпеки з підтримкою ШІ.

Впровадження зон фокусування (AoF) підкреслює важливість конфіденційності в механізмах зворотного зв'язку, що використовуються для звітів про кібератаки та спільних баз даних в аналізі ризиків кіберфізичних систем (CPS). Переваги використання Штучного Інтелекту в кібербезпеці в тому, що це покращує точність та швидкість виявлення кіберзагроз,а також автоматизація рутинних завдань, прогнозування та управління кіберризиками, та зниження витрат на кібербезпеку.

Інтеграція штучного інтелекту в кіберфізичні системи призвела до швидкої появи досліджень і порівнянь в літературі, яка змінила не тільки аналітику кіберризиків, але й аналітику даних. У цій роботі представлена нова структура, яка пояснює, як ШІ можна інтегрувати з аналітикою кіберризиків. Це підтверджує, що розробка CPS вимагає розуміння проектування системи, системної інженерії та соціології системи.

Основні висновки включають:

1.Інтеграція штучного інтелекту в комунікаційні мережі та підключену технологію має розвиватися в етичному порядку, зрозумілому людям, зберігаючи при цьому максимальну довіру та конфіденційність користувачів;

2.Координація штучного інтелекту в CPS має бути надійною, щоб запобігти зловживанню з боку внутрішніх загроз, організованої злочинності, терористичних організацій або спонсорованих державою агресорів;

3.Ризик даних спонукає приватний сектор вживати заходів для покращення управління інтелектуальною власністю конфіденційної та закритої інформації та захисту інформації, що дозволяє ідентифікувати особу;

4.Аналіз динамічного та самоприйнятного дизайну штучного інтелекту для механізму когнітивної системи для контролю, аналізу, розподілу та керування імовірнісними даними.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. John Maynard Keynes, «Economic Possibilities for our Grandchildren» (1930), in Essays in Persuasion (New York: Harcourt Brace, 1932), 358–373. URL: https://www.aspeninstitute.org/wpcontent/uploads/files/content/upload/Intro_and_Section_I.pdf

2. Elon Musk Reminds Us of the Possible Dangers of Unregulated AI. URL: https://futurism.com/elon-musk-reminds-us-of-the-possible-dangers-of-unregulated-ai

3. Стівен Хокінг: штучний інтелект може стати найгіршим винаходом людства. URL: https://mind.ua/news/20178313-stiven-hoking- shtuchnij-intelekt-mozhe-stati-najgirshim-vinahodom-lyudstva.

4. Штучний інтелект стане головною загрозою людству. URL: https://www.bbc.com/ukrainian/features-65728291

5. Facebook вимкнув систему штучного інтелекту: боти винайшли свою мову. URL: https://www.unian.ua/science/2058576-facebook-vimknuv-sistemu-shtuchnogo-intelektu-boti-vinayshli-svoyu-movu.html

6. Восстание машин. Как искусственный интеллект родил цифрового ребёнка. URL: https://www.dsnews.ua/future/vosstanie-mashin-kak-iskusstvennyy-intellekt-rodil-tsifrovogo-12122017220000

7. Джон Маркофф. Homo Roboticus? Люди и машины в поисках взаимопонимания. URL: http://loveread.ec/read_book.php?id=64814&p=1

8. Єфремов М. Штучний інтелект, історія та перспективи розвитку / М. Єфремов. URL: https://naub.oa.edu.ua/shtuchnyj-intelekt-istoriya-ta-perspektyvy/

9. Четверта промислова революція заради Землі Використання можливостей штучного інтелекту на користь Землі. Січень 2018 р. URL: https://www.pwc.com/ua/uk/survey/2018/ai-for-the-earth-jan-2018.pdf

10. Штучний інтелект як технологія створення автоматизованих інтелектуальних систем. URL: https://er.knutd.edu.ua/bitstream/123456789/5044/1/20160428-29_TEZY_V3_P349.pdf

11. Поняття штучного інтелекту. URL: http://megalib.com.ua/content/1956_71_P

onyattya_shtychnogo_intelekty.html

12. Штучний інтелект. Підходи і напрямки до розуміння штучного інтелекту. URL: http://referat-ok.com.ua/work/shtuchnij-intelekt/

13. Переваги та недоліки застосування штучного інтелекту у сферах управління. URL: http://elartu.tntu.edu.ua/bitstream/lib/25207/2/MSNK_2018v2_Pelcher_M-Advantages_and_lack_ of_application_72-73.pdf

14. Moshe Y. Vardi, «The Consequences of Machine Intelligence». Atlantic, October 25, 2012. URL: http://www.theatlantic.com/ technology/archive/2012/10/the-consequences-of-machine- intelligence/264066

15. Штучний інтелект: що це і яку несе небезпеку. URL: https://24tv.ua/techno/shtuchniy_intelekt_shho_tse_i_yaku_ nese_nebezpeku_n914662

16. John McCarthy, book review of B.P. Bloomfield, The Question of Artificial Intelligence: Philosophical and Sociological Perspectives, in Annals of the History of Computing 10, no. 3 (1988): 224–229

17. Улянівський Т. Штучний інтелект – це продовження еволюції. URL: https://zbruc.eu/node/71907

18. Області практичного застосування систем штучного інтелекту. URL: https://sites.google.com/site/eksperntisistemi/ zastosuvanna-sistem-stucnogo-intelektu

19. Панченко В., Резнікова Н. Повстання машин. Чи замкне штучний інтелект коло фінансового зубожіння. URL: http://dniprograd.org/2017/08/31/povstannya-mashin-chi- zamkne-shtuchniy-intelekt-kolo-finansovogo-zubozhinnya_59965.

20. Мартін Форд. Пришестя роботів: техніка і загроза майбутнього безробіття / М. Форд. URL: http://www.e-reading.club/ bookreader.php/1057296/Ford_-_Prishestya_robotiv_tehnika_i_ zagroza_maybutnogo_bezrobittya.html.

21. Співзасновник «Google» розповів, яку загрозу несе штучний інтелект. URL : https://ua.korrespondent.net/tech/science/ 3966047-spivzasnovnyk-Google-rozpoviv-yaku-zahrozu-nese- shtuchnyi-intelekt.

22. Ник Бостром. Штучний інтеллект. Етапи.Загрози.Наслідки. URL: http://loveread.ec/view_global.php?id=70922.

23. MIT is spending $1 billion to open a college in 2019 just for AI. URL: https://news.mit.edu/2018/mit-reshapes-itself-stephen-schwarzman-college-of-computing-1015.

24. M.I.T. Plans College for Artificial Intelligence, Backed by $1 Billion. URL: https://www.nytimes.com/2018/10/15/technology/ mit-college-artificial-intelligence.html.

25. AI Forum of New Zealand і AsureQuality, «Штучний інтелект для сільського господарства в Новій Зеландії», с. 40, 2019. URL: https://aiforum.org.nz/wp-content/uploads/2019/10/Artifical-Intelligence-For-Agriculture-in-New-Zealand.pdf

26. Д. Ву та інші, «Кібербезпека для цифрового виробництва», J. Manuf. сист., вип. 48, стор. 3–12, 2018 URL: https://www.sciencedirect.com/science/article/abs/pii/S0278612518300396.

27. Штучний інтелект (AI): Що це таке і чому це важливо? URL: http://surl.li/ezkrn

28. Штучний інтелект навчився діагностувати очні хвороби. URL: http://bukovina.biz.ua/news/48130

29. T. C. Truong, Q. B. Diep, and I. Zelinka, "Artificial intelligence in the cyber domain: Offense and defense,"*Symmetry (Basel).*, vol. 12, no. 3, pp. 1–24, 2020 URL: https://www.mdpi.com/2073-8994/12/3/410

30. I. H. Sarker, Y. B. Abushark, F. Alsolami, and A. I. Khan, "IntruDTree: A machine learning based cyber security intrusion detection model," *Symmetry (Basel).*, vol. 12, no. 5, pp. 1–15, 2020 URL: https://www.mdpi.com/2073-8994/12/5/754

31. S. Bredt, "Artificial Intelligence (AI) in the Financial Sector—Potential and Public Strategies," *Front. Artif. Intell.*, vol. 2, no. October, pp. 1–5, 2019 URL: https://www.frontiersin.org/articles/10.3389/frai.2019.00016/full

32. Z. Siddiqui, M. S. Husain, and S. Yadav, "Application of Artificial Intelligence in Fighting Against Cyber Crimes: a Review," *Int. J. Adv. Res. Comput. Sci.*, vol. 9, no. 2, pp. 118–122, 2018

33. A. Ren, D. Wu, W. Zhang, J. Terpenny, and P. Liu, "Cyber security in smart

manufacturing: Survey and challenges," *67th Annu. Conf. Expo Inst. Ind. Eng. 2017*, pp. 716–721, 2017.

34. M. N. O. Sadiku, O. I. Fagbohungbe, and S. M. Musa, "Artificial Intelligence in Cyber Security," *Int. J. Eng. Res. Adv. Technol.*, vol. 06, no. 05, pp. 01–07, 2020, URL: https://ijerat.com/index.php/ijerat/article/view/424/423

35. "The Web Application Hacker's Handbook: Finding and Exploiting Security Flaws" Dafydd Stuttard Marcus Pinto URL: https://books.google.com.ua/books?id=jN6cDprdnd0C&printsec=frontcover&redir_esc=y#v=onepage&q&f=false

36. Intrusion Detection Systems by Rebecca Bace and Peter Mell URL: http://cs.uccs.edu/~cchow/pub/ids/NISTsp800-31.pdf

37. Comprehensive Review of Malware Detection Techniques November 2021 Conference: 2021 International Conference on Innovative Computing (ICIC)

38. IoT-Enabled Healthcare: Benefits, Issues and Challenges November 2020 Conference: THE 4TH INTERNATIONAL CONFERENCE ON FUTURE NETWORKS AND DISTRIBUTED SYSTEMS URL: https://www.researchgate.net/publication/351391253_IoT-Enabled_Healthcare_Benefits_Issues_and_Challenges

39. Aleroud A., Zhou L. Phishing environments, techniques, and countermeasures: a survey. Computers & security. 2017. Vol. 68. P. 160–196. URL: https://doi.org/10.1016/j.cose.2017.04.006

40. Тенденції штучного інтелекту в кібербезпеці, на які варто звернути увагу в 2024 році URL: http://surl.li/sgrsd

41. Hariri RH, Fredericks EM, Bowers KM (2019) Uncertainty in big data analytics: survey, opportunities, and challenges. J Big Data URL: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0206-3

42. De Roure D, Page KR, Radanliev P, Van Kleek M (2019) Complex coupling in cyber-physical systems and the threats of fake data. In: Living in the internet of things (IoT 2019), URL:

https://www.researchgate.net/publication/336341104_Complex_coupling_in_cyber-physical_systems_and_the_threats_of_fake_data

43. Hassanien, A., Haqiq, A., Tonellato, P., Bellatreche, L., Goundar, S., & Azar, A. et al. Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2021).

44. Keen, E. (2021). The benefits and limitations of AI in cybersecurity - Help Net Security. Help Net Security. Retrieved 1 September 2021, URL: https://www.helpnetsecurity.com/2018/12/20/ai-cybersecurity-benefits-limitations/.

45. Raghavan, V., Venkat N. Gudivada, & Venu Govindaraju. (2016). Cognitive Computing: Theory and Applications. Elsevier Science. URL: https://vdoc.pub/documents/cognitive-computing-theory-and-applications-1q4nl2u8c778

46. What are the AI security risks? URL: https://www.tarlogic.com/blog/ai-security-risks/

47. Ethics of Artificial Intelligence: Case Studies and Options for Addressing Ethical Challenges URL: https://www.researchgate.net/publication/366764402_Ethics_of_Artificial_Intelligence_Case_Studies_and_Options_for_Addressing_Ethical_Challenges

48. ISO 31000:2018 Risk management — Guidelines https://www.iso.org/obp/ui/#iso:std:iso:31000:ed-2:v1:en