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

Кафедра кібербезпеки

КВАЛІФІКАЦІЙНА РОБОТА

на тему:

«ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ У СФЕРІ КІБЕРБЕЗПЕКИ»

Виконала: студентка 4 курсу
Рівень бакалавр
Галузь знань 12 «Інформаційні технології»,
спеціальність 125 «Кібербезпека»
Каракян Афіна Володимирівна

Керівник: к.тех.н., доцент
Ахмаметьєва Ганна Валеріївна

Рецензент: д.ф.м.н. професор
Василенко Микола Дмитрович

Одеса – 2024

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

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

З А В Д А Н Н Я
на кваліфікаційну (бакалаврську) роботу
здобувачу Каракян Афіні Володимирівні

1. Тема роботи: «Використання штучного інтелекту у сфері кібербезпеки»
Керівник роботи: к.т.н, доцент Ахмаметьєва Ганна Валеріївна
затверджена наказом НУ ОЮА № 809-06 від «02» квітня 2024 року
2. Строк подання здобувачем роботи «20» травня 2024 року
3. Дата видачі завдання«15» вересня 2023 року

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів бакалаврської роботи	Строк виконання етапів роботи	Примітка

Здобувач _____
(підпис) (прізвище та ініціали)

Керівник роботи _____
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Кваліфікаційна робота на тему "Використання штучного інтелекту у сфері кібербезпеки" на здобуття першого (бакалаврського) рівня вищої освіти за спеціальністю 125 Кібербезпека, освітньою програмою: Кібербезпека, містить 10 рисунків, 16 літературних джерел за переліком посилань. Робота виконана на 33 сторінках загального тексту і 29 сторінках основного тексту.

Метою даної кваліфікаційної роботи є аналіз актуальності застосування штучного інтелекту у сфері кібербезпеки та його впливу на забезпечення безпеки цифрового простору в найближчому майбутньому.

Під час роботи було розглянуто різні варіанти застосування штучного інтелекту у сфері кібербезпеки, можливий розвиток штучного інтелекту у майбутньому., а також було проведено розгляд різних типів захисту за допомогою штучного інтелекту.

Практична значеність цієї роботи полягає у зборі та аналізі найперспективніших сучасних варіантів розвитку та навчання штучного інтелекту для подальшого використання у сфері кібербезпеки.

CYBERSECURITY, ШТУЧНИЙ ІНТЕЛЕКТ, ІНФОРМАЦІЙНА БЕЗПЕКА, МАШИННЕ НАВЧАННЯ

ANNOTATION

The qualification work on the topic "The Use of Artificial Intelligence in the Field of Cybersecurity" for obtaining the first (bachelor's) level of higher education in the specialty 125 Cybersecurity, educational program: Cybersecurity, contains 10 figures, 16 literature sources according to the list of references. The work is completed on 33 pages of total text and 29 pages of the main text.

The purpose of this qualification work is to analyze the relevance of the use of artificial intelligence in the field of cybersecurity and its impact on the security of the digital space in the near future.

During the work, various options for the use of artificial intelligence in the field of cybersecurity, the possible development of artificial intelligence in the future were considered, as well as various types of protection using artificial intelligence.

The practical significance of this work is to collect and analyze the most promising modern options for the development and training of artificial intelligence for further use in the field of cybersecurity.

CYBERSECURITY, ARTIFICIAL INTELLIGENCE, INFORMATION SECURITY, MACHINE LEARNING

ЗМІСТ

ВСТУП

Так як у наш час штучний інтелект розвивається стрімкіше, ніж будь-коли, у нього з'являється все більше сфер застосування, а те, що раніше було примітивним дослідженням - зараз має неймовірні перспективи розвитку. Безліч передових компаній по всьому світу зацікавлені в застосуванні штучного інтелекту для виконання професійних завдань, зокрема тих, які часто потребують тривалого часу на виконання вручну. Штучний інтелект на відміну від будь-яких інших написаних людиною алгоритмів володіє власним координованим мисленням і має можливість навчатися самостійно на основі отриманого матеріалу.

У зв'язку з тим, що сучасний світ переживає справжню революцію в галузі технологій, штучний інтелект стає не лише передовим інструментом, але й ключовим стратегічним ресурсом у боротьбі з кіберзагрозами. Його потужний аналітичний потенціал та здатність до навчання на основі великих обсягів даних робить його важливим активом у виявленні та протидії різноманітним кібератакам.

Процеси автоматизації, підтримані штучним інтелектом, дозволяють удосконалити моніторинг та аналіз кіберзагроз, що в свою чергу дозволяє швидше реагувати на нові загрози та атаки. Але разом із зростанням потужності штучного інтелекту з'являються нові етичні, правові та соціальні виклики, такі як приватність даних, алгоритмічне відхилення і відповідальність за прийняті рішення. Усвідомлення цих проблем дозволяє глибше розуміти контекст використання штучного інтелекту в кібербезпеці та його можливі наслідки.

Зокрема, потрібно акцентувати увагу на розвитку ефективних механізмів регулювання та контролю за використанням штучного інтелекту в цілях забезпечення безпеки та захисту особистих даних. Також важливо розглядати перспективи майбутнього розвитку штучного інтелекту та його потенціал для розв'язання складних проблем у кібербезпеці, таких як автоматизація виявлення вразливостей та прогнозування майбутніх атак.

На фоні стрімкого розвитку технологій, штучний інтелект стає ключовим фактором в сфері кібербезпеки. Нинішній стан справ свідчить про необхідність

вивчення впливу штучного інтелекту на безпеку кіберпростору. Незважаючи на досягнення у цій області, виникають нові виклики та загрози, що вимагають системного підходу та розробки ефективних стратегій захисту.

Метою даної кваліфікаційної роботи є аналіз актуальності застосування штучного інтелекту у сфері кібербезпеки та його впливу на забезпечення безпеки цифрового простору в найближчому майбутньому.

Потрібно виконати наступні задачі:

1. Провести аналіз поточного стану кібербезпеки та основних загроз у цифровому просторі.
2. Дослідити методи застосування штучного інтелекту у сфері кібербезпеки.
3. Розглянути переваги та недоліки використання штучного інтелекту в кібербезпеці та визначити потенційні ризики та виклики, пов'язані із застосуванням штучного інтелекту для забезпечення безпеки цифрового простору.
4. Проаналізувати перспективи розвитку технологій штучного інтелекту у сфері кіберзахисту.
5. Оцінити вплив штучного інтелекту на забезпечення кібербезпеки в найближчому майбутньому.

1 ТЕОРЕТИЧНІ АСПЕКТИ ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ У КІБЕРБЕЗПЕЦІ

1.1 Загальні відомості про штучний інтелект та його розвиток

Штучний інтелект (ШІ) являє собою систему, здатну навчатися самостійно і взаємодіяти з користувачем, використовуючи згенеровані нею самою патерни з урахуванням наданих обмежень і вказівок. Він здатний аналізувати величезні обсяги даних у реальному часі набагато швидше за людину та надавати дані за будь-яким критерієм.

До недавнього часу штучний інтелект був вкрай примітивним, його здібності були побудовані на звичайних алгоритмах виключення варіантів. але машинне навчання нейромереж дало значний потенціал розвитку всієї сфери ШІ.

Наразі нейромережі здатні генерувати зображення, звук, імітувати голос людини, писати код, частково використовувати критичне мислення, за малу кількість часу опрацьовувати велику кількість даних, миттєво знаходити рішення будь-яких запитів користувача, а в професійній сфері навіть замінювати людину у виконанні великої кількості монотонної роботи та надавати точні аналізи ґрунтуючись на отриманій інформації.

У сфері кібербезпеки штучний інтелект відіграє велику роль. Якщо раніше його використовували лише для відстеження мережевого трафіку за встановленими правилами, то тепер же він самостійно здатний виокремлювати алгоритми для прогнозування кібератак, обробляючи інформацію та вивчаючи всі послідовності їхнього виникнення для подальшої ідентифікації схожих ознак.

Штучний інтелект в сфері кібербезпеки відіграє рішучу роль, надаючи можливість не лише ефективно виявляти аномалії та загрози, а й передбачати їхні майбутні вектори розвитку. Завдяки розширеним можливостям аналізу даних та розумінню складних зв'язків, штучний інтелект дозволяє реалізувати відповідальний та продуманий підхід до кіберзахисту.

Крім того, автоматизація процесів виявлення та реагування на кіберзагрози за допомогою штучного інтелекту дозволяє значно збільшити швидкість реакції та

зменшити час, необхідний для виявлення та вирішення потенційних проблем. Це важливо не лише для забезпечення безпеки та стабільності цифрового простору, а й для збереження репутації та довіри користувачів.

Це дозволяє звільнити людей від рутинних завдань та сконцентрувати їх на стратегічних аспектах кібербезпеки, таких як розробка та впровадження нових стратегій захисту, аналіз вразливостей та розробка активних заходів протидії потенційним загрозам. Такий підхід сприяє збільшенню ефективності та швидкості реагування на кібератаки, що є критичним у сучасному цифровому середовищі.

1.2 Роль штучного інтелекту в сучасних системах кібербезпеки

Штучний інтелект як галузь науки виник у 1956 році завдяки роботі чотирьох американських вчених - Шеннона, Карті, Рочестера і Мінського - на семінарі в Дартмутському коледжі в США. Між 1910 і 1913 роками Бартранд Рассел і Вайтхед опублікували «Принципи математики» (Principia Mathematica), представивши нову ідею поєднання математики і логіки. Конрад Цузе спроектував перший комп'ютер з програмним управлінням у 1941 році. Термін «нейронна мережа» з'явився в остаточному вигляді в «Логічному обчисленні ідей, іманентних нервовій діяльності» Уоррена Макклака і Волтера Піттса.

Починаючи з 1980-х років, штучний інтелект почали впроваджувати в кібербезпеку. Перші спроби полягали у впровадженні систем оповіщення, які слідували за спеціально розробленими параметрами для запуску сповіщень. Потім, через два десятиліття, прогрес у машинному навчанні у 2000-х роках допоміг штучному інтелекту ідентифікувати типові моделі трафіку та поведінку користувачів на підприємстві, що дозволило швидко реагувати на незвичну та ризиковану активність.

Нинішня робота штучного інтелекту для кібербезпеки схожа на принцип, описаний вище, що застосовувався у 2000-х. Головне його завдання - опрацювання великих обсягів даних із кількох джерел, контроль трафіку і визначення його закономірності під час використання в різних організаціях. Сюди входить

опрацювання таких даних як: час і місце розташування під час входу в систему, аналіз пристроїв і трафіку пристроїв, з яких співробітники входили в систему. З часом штучний інтелект навчається розпізнавати типову для користувача поведінку, надалі відрізняючи її від аномальної.

Важливо зазначити, що дані організації не використовуються для передачі даних штучному інтелекту інших організацій, щоб забезпечити захист даних від несанкціонованого доступу. Використовуючи ці два фактори, штучний інтелект використовує глобальну інформацію про загрози, створену шляхом об'єднання інформації з декількох організацій.

Наразі існують як суттєві за, так і проти введення штучного інтелекту, одним із таких побоювань поділився Джефрі Хітсон, учений у галузі інформатики. У газеті The New York Times, випущеній у травні 2023 року, було опубліковано заяву про звільнення ним посади віцепрезидента компанії Google. Ним було уточнено, що такими діями він хоче попередити суспільство про небезпеки та ризики, пов'язані з впровадженням технології штучного інтелекту в їхній продукт.

Більшість його тез є суто теоретичними, і вченого дуже турбує перспектива розвитку штучного інтелекту, який неможливо зупинити без належного контролю або застосування суворих обмежень. До 2022 року, коли великі компанії на кшталт Google та OpenAI починають створювати власні моделі штучного інтелекту, що подекуди перевершують людський, позиція дослідника починає кардинально змінюватися. З ним погоджуються багато впливових вчених: Ілон Маск, Стів Возняк та понад 1000 інших експертів закликали до шестимісячного мораторію на навчання ШІ, щоб вирішити питання етики та безпеки.

1.3 Переваги та недоліки використання штучного інтелекту в кібербезпеці

Значною перевагою штучного інтелекту є його гнучкість та адаптивність. Штучний інтелект може швидко адаптуватися до нових загроз та змінюватися відповідно до поточної ситуації. Він може аналізувати великі обсяги даних у реальному часі та швидко реагувати на нові вектори атак або виявлені вразливості.

Також однією з важливих переваг штучного інтелекту є його здатність аналізувати та розуміти складні патерни в кіберзлочинах, які можуть залишитися непоміченими людськими аналітиками. Використання алгоритмів машинного навчання дозволяє виявляти та розпізнавати навіть найбільш витончені методи атак.

Проблема використання штучного інтелекту полягає у зловживанні ним шахраями. У наш час дедалі більше зростає кількість шахрайських схем з використанням штучного інтелекту для підроблення голосу або особистості у цілому. Часто такі злочини пов'язані з вимаганням коштів.

Другою проблемою також можна вказати питання приватності штучного інтелекту. Обробка великих обсягів даних може спричинити загрозу безпеці для особистих даних і конфіденційності окремих осіб.

З цього питання випливає наступне: питання етичності. Для штучного інтелекту потрібні чіткі обмеження під час використання та моніторингу даних для позначення того, який рівень втручання є прийнятним, а який порушує конфіденційність користувача.

Також варто відзначити можливу залежність від штучного інтелекту. Нестача людського контролю та уважність може нашкодити сучасній кібербезпеці, оскільки штучний інтелект не є прямим виконавцем, а лише помічником, у той час як безпека людини у кіберпросторі насправді залежить від неї самої.

Окрім цього через штучний інтелект може виникнути нестача кваліфікованих фахівців. Цей пункт має дві сторони: з одного боку, штучний інтелект може на певний час замінити потрібних співробітників, але це виключно тимчасове рішення.

Штучний інтелект має обмеження в творчому мисленні. Хоча ШІ демонструє високі результати у вирішенні лінійних завдань, він стикається зі значними труднощами при вирішенні творчих задач та завдань, що вимагають інтуїтивного підходу. Наприклад, системи штучного інтелекту не здатні генерувати абсолютно нові ідеї або відчувати емоції, що є важливими складовими творчого процесу. Це

суттєво обмежує його застосування в галузях, де необхідна креативність та інноваційність.

Крім того, однією з найбільш значущих причин обмеженого впровадження ШІ є його висока вартість. Процес розробки та впровадження штучного інтелекту потребує значних фінансових інвестицій. Вартість включає створення самих систем, їх навчання на великих обсягах даних, а також постійне вдосконалення та адаптацію до нових умов. Ці витрати можуть бути непосильними для багатьох організацій.

Важливо також зазначити, що існує суттєвий розрив у безпеці між країнами з різним рівнем розвитку технологій штучного інтелекту. Це створює додаткові ризики та вразливості, оскільки країни з меншими фінансовими можливостями не можуть забезпечити належний рівень захисту своїх систем. Така нерівність може призвести до збільшення кіберзагроз та глобальної нестабільності.

2 РОЗВИТОК ШТУЧНОГО ІНТЕЛЕКТУ ТА ЙОГО ВПЛИВ НА КІБЕРБЕЗПЕКУ

2.1 Виклики та ризики, пов'язані зі зростанням використання штучного інтелекту в кіберпросторі

Використання штучного інтелекту в сучасному світі стикається з безліччю ризиків, які так чи інакше стосуються безпеки. В історії та індустрії було безліч прикладів, коли люди намагалися створити досконалу машину, що виконує чіткі вимоги, яка не виходить за рамки свого коду і здатна вчасно зупинитися, але майже в кожному випадку усе закінчувалося провалом. (див. рис. 2.1)

Рисунок 2.1 - Приклад невдалої розробки Microsoft штучного інтелекту з доступом до відкритого спілкування з людьми у Twitter

Саме тому при створенні такого складного програмного засобу використовуються не тільки фундаментальні поняття і закони, здатні працювати відповідно до класичного тесту Тюрінга, а й сучасні протоколи безпеки, які в разі виявлення збою або неполадки готові негайно зупинити процес подальшої роботи або сповільнити його для подальшого втручання зі сторони.

Сучасна розробка штучного інтелекту має зосереджуватися на наступних пунктах:

- Безпека робочого простору користувача

- Алгоритмічна нерівність і справедливість
- «Прозорість» чинної системи
- Нормативно-правові та психологічні аспекти користувача

Безпека робочого простору користувача охоплює заходи забезпечення безпеки та захисту інформації у робочому середовищі користувача. Штучний інтелект може використовуватися для виявлення та запобігання кіберзагрозам, таким як віруси, шкідливі програми або хакерські атаки. Наприклад, системи машинного навчання можуть аналізувати мережевий трафік та виявляти аномальні активності, що можуть вказувати на потенційні загрози для безпеки.

Алгоритмічна нерівність і справедливість у більшості стосується етичних аспектів використання штучного інтелекту. Наприклад, в алгоритмах прийняття рішень може виявитися алгоритмічна нерівність, коли певні групи людей або певні категорії даних мають перевагу або недоліки. Розробка справедливих алгоритмів та врахування різних соціальних та культурних контекстів є важливими аспектами у забезпеченні етичного використання штучного інтелекту.

«Прозорість» чинної системи відноситься до зрозумілості та інтерпретованості рішень, прийнятих штучним інтелектом. Штучний інтелект може приймати складні рішення на основі великих обсягів даних та складних алгоритмів, але важливо, щоб ці рішення були зрозумілими та перевірені людськими експертами. Наприклад, системи машинного навчання можуть бути навчені розпізнавати об'єкти на зображеннях, але важливо знати, як саме вони приймають ці рішення та які чинники впливають на їхню точність.

Пункт про нормативно-правові та психологічні аспекти користувача охоплює важливість розробки законодавства та правил використання штучного інтелекту з урахуванням прав користувачів та їх психологічних аспектів. Наприклад, важливо враховувати питання приватності та захисту даних у контексті використання штучного інтелекту, а також розглядати питання впливу штучного інтелекту на психічний стан користувачів та їхні соціальні відносини.

Безумовно, крім кіберпростору, у штучного інтелекту також є й інші важливі питання, що стосуються етики, майбутнього і «почуттів» користувача, але це вже

більш масштабні питання, які потребують розв'язання поза технологіями і взаємодією людини з ними. Розглянемо ж кожен із вищенаведених пунктів докладніше на прикладах.

Напевно, найвідоміша на даний момент проблема у використанні штучного інтелекту полягає в тому, щоб надійно захистити кіберпростір, у якому працює користувач. Питання, пов'язані з конфіденційністю, захист даних від паразитуючих програм, вірусів і хакерських атак, безпека використання отриманої інформації - всі ці аспекти потрібно враховувати.

2.2 Перспективні напрямки досліджень та інновації у сфері застосування штучного інтелекту для захисту інформаційних систем

Сьогодні розробка штучного інтелекту є основною метою багатьох дослідницьких проектів, з метою досягнення сильного штучного інтелекту до 2020 року. На відміну від людського розуму, на розвиток якого пішли мільйони років, ШІ може розвиватися в геометричній прогресії, засвоюючи величезні обсяги нової інформації за лічені секунди.. (див. рис. 2.2)

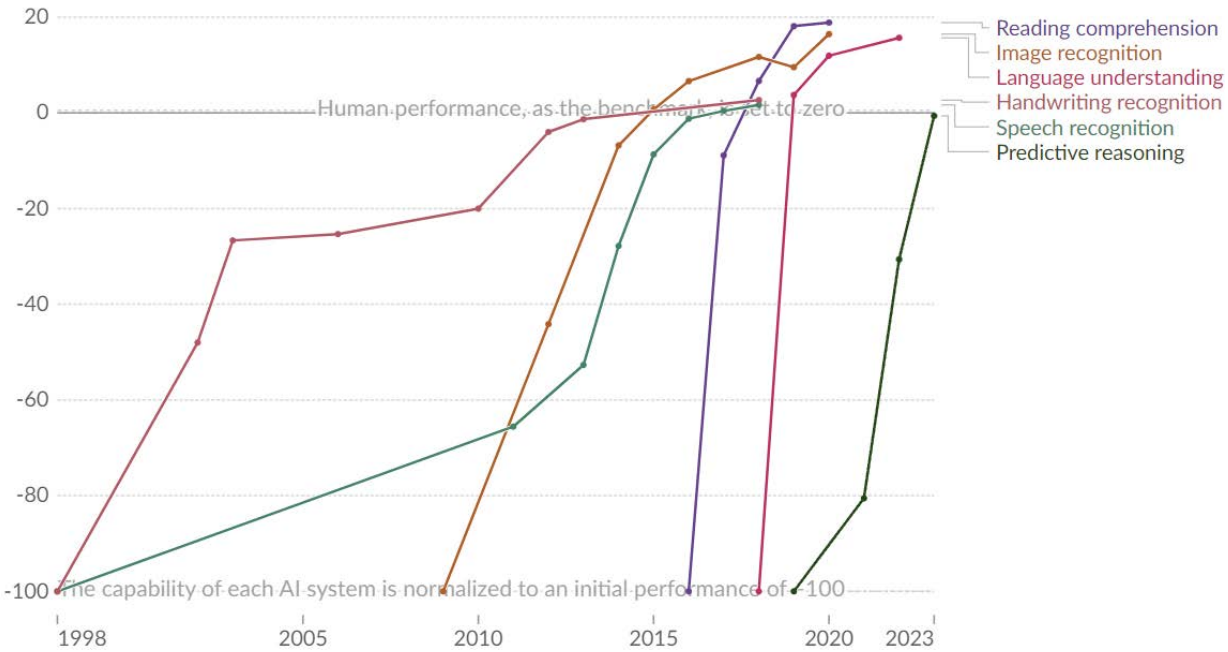

Рисунок 2.2 - Результати тестування систем штучного інтелекту на різні можливості порівняно з людською продуктивністю

З цих причин дослідження штучного інтелекту наразі проводяться більшістю керівних компаній у світі. Google починаючи з 2010 року викуповував і адаптував велику кількість проектів і досліджень, які пізніше ставали частиною Google Brain та Google Deep Mind і наразі є однією з найперспективніших компаній, що розробляють AGI.

Компанія Microsoft також оприлюднювала матеріал власних тестувань системи GPT-4, виявивши в ній ознаки близькі до сильного штучного інтелекту. Наприклад, незважаючи на те, що GPT-4 навчався тільки на текстових даних, він самостійно виробив навичку створення зображень. (див. рис. 2.3)

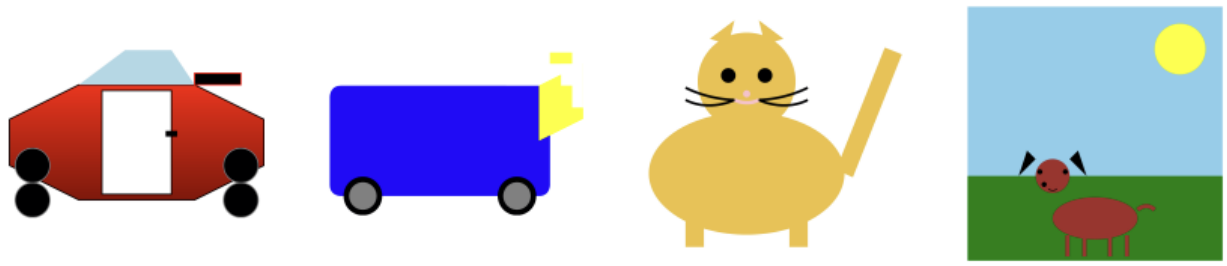

Рисунок 2.3 - Зразки згенерованих SVG зображень з матеріалу оприлюдненого компанією Microsoft

На думку різних фахівців, перспектива глобального розвитку штучного інтелекту досягне свого піку вже в 2029 році або навіть раніше.

Розвиток систем автоматизованого реагування на інциденти (API) стає ключовим напрямком в області кібербезпеки. Штучний інтелект може значно полегшити процес виявлення, аналізу та реагування на кібератаки, забезпечуючи швидку та ефективну відповідь на потенційні загрози.

Загалом, розвиток штучного інтелекту в сфері кібербезпеки відкриває широкі можливості для створення нових інноваційних рішень, що дозволять забезпечити надійний захист інформаційних систем в умовах непередбачуваних кіберзагроз.

2.3 Тест Тюрінга, ключові відмінності штучного інтелекту від людини та Captcha

Тест Тюрінга - тест, названий на честь Алана Тюрінга. Початкова ідея цього тесту була сформована в статті під назвою «Обчислювальні машини і розум» (англ. Computing Machinery and Intelligence), яка побачила світ у 1950 році. Історія має безліч підстав для проведення цього тесту і виникнення питань про здібності інтелекту різних «розумних істот».

Метою цього дослідження є визначити, чи здатна машина виявляти ознаки інтелекту. Оригінальний тест передбачає наявність 3 учасників: комп' ютерної програми, людини та судді. Останній, без участі у спілкуванні опонентів, мав визначити хто є машиною, а хто - людиною, використовуючи винятково текстові повідомлення, надані ними обома. Текстовий формат проведення даного експерименту був обраний для чистоти втілення, адже спілкування усним мовленням краще покаже здатність програми сприймати та відтворювати його, ніж сконцентруватися на інтелекті. На листування відводиться певний час для неможливості формування відповіді за її швидкістю. Це правило не застаріло, оскільки раніше комп'ютери реагували набагато повільніше за людину, зараз же вони навпаки здатні робити це набагато швидше. У цьому тесті машина вважається переможцем, якщо суддя не може точно сказати, чий текст він бачить перед собою.

Механізм роботи цього тесту певною мірою схожий з механізмом Captcha. Можна сказати, що Captcha є його інтерпретацією для сучасної реальності. Перед тим, як потрапити на сайт і надалі виконувати на ньому якісь дії, користувачеві надають певний набір зображень з об'єктами на них, зображення з викривленим текстом або аудіо-версію цього тесту, що являє собою спотворений запис. Для проходження цього тесту від користувача потребується вказати на зображення із правильним набором об'єктів або введення правильного тексту у відповідне поле.

3 ПРАКТИЧНЕ ЗАСТОСУВАННЯ ШТУЧНОГО ІНТЕЛЕКТУ У КІБЕРБЕЗПЕЦІ

3.1 Машинне навчання та його застосування у виявленні загроз

Машинне навчання є основою для автоматизації, запам'ятовування різних шаблонів дій. Для цього існує три основних способи машинного навчання.

1. Навчання з додатковою допомогою від людей.
2. Самостійне навчання.
3. Самостійне навчання з допомогою додаткової інформації.

Навчання з додатковою допомогою від людей означає, що штучний інтелект буде навчатися як школяр із репетитором. Їй дають задачі та відповіді для тренування, а вона аналізує дані, зіставляє свої висновки з правильними рішеннями та налаштовує себе, щоб ставати точнішою. (див. рис. 3.1)

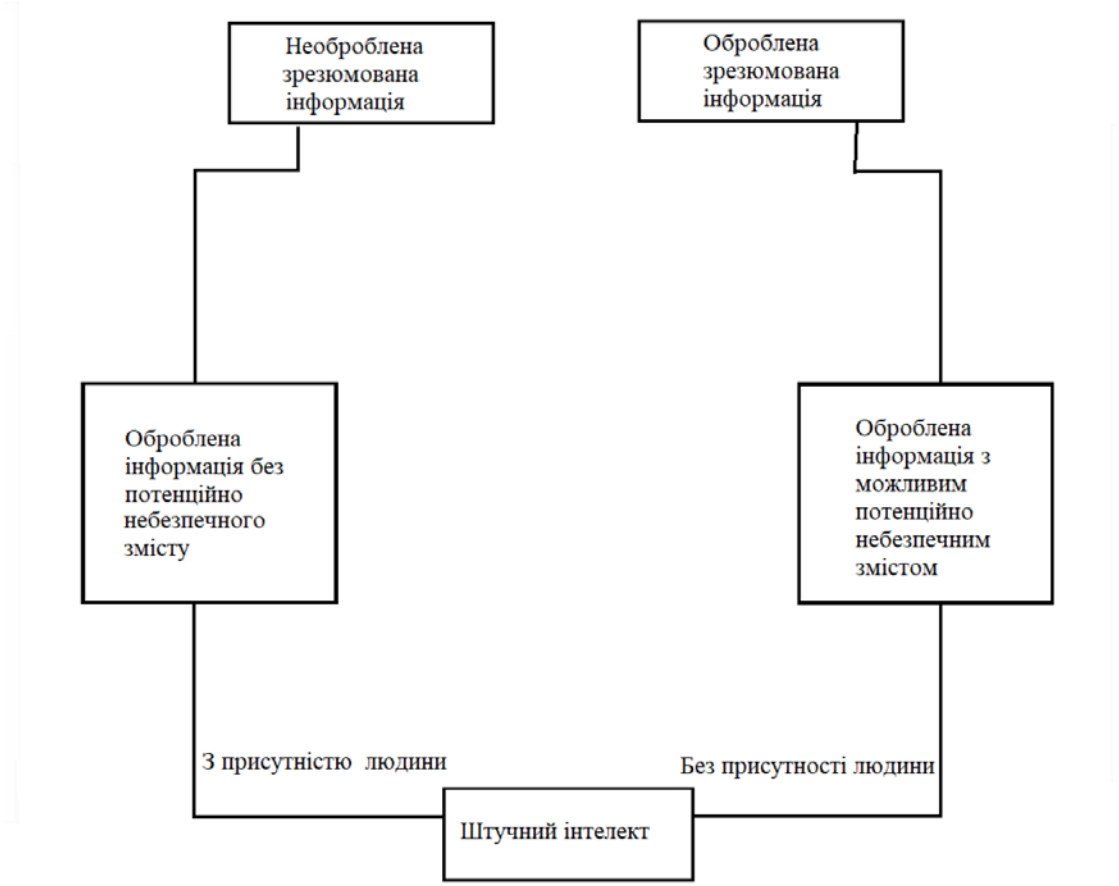

Рисунок 3.1 – Приклад схеми навчання найпростішого штучного інтелекту з додатковою допомогою від людей

Це допоможе штучному інтелекту навчитись швидше надавати відповіді/результат після отримання задачі. У результаті подібного навчання модель буде здатною самостійно виконувати нові завдання, що будуть їй поставлені та аналізувати усе спираючись на отриманий досвід під час навчання.

Цей метод добре підходить, коли треба класифікувати задачі за різними типами. Наприклад, штучний інтелект такого типу здатний виявити проблему у тексті за словником або зробити висновок стосовно відношення файла до вірусних програм за допомогою шаблонів на яких його тренували. Через те що саме людина є основною частиною цієї моделі навчання, вона використовується лише у специфічних випадках.

Самостійне навчання є моделлю машинного навчання, яка працює з нерозміченими даними, тобто без відповідей. Можна сказати модель навчається на власних помилках. Її задача – знайти приховані структури й закономірності в них.

За допомогою такого типу машинного навчання можна групувати об’єкти зі схожими характеристиками, виявляти аномалії, коли щось не вписується в очікувані шаблони (підозрілі активності в мережі, нові пристрої, які не зареєстровані в системі).

Машинне навчання без вчителя особливо корисне, коли даних багато і складно побачити між ними зв’язок (Користувачу невідомо заздалегідь на які категорії їх ділити). (див. рис. 3.2)

Самостійне навчання з допомогою додаткової інформації. У цьому випадку система вчиться шляхом взаємодії з навколишнім середовищем, отримуючи нагороди за успіхи й покарання за помилки. Її мета – зібрати якомога більше нагород, тобто знайти найкращий спосіб діяти. Система отримує задачу та інформаційні ресурси, які допоможуть із виконанням цієї задачі. (див. рис. 3.3)

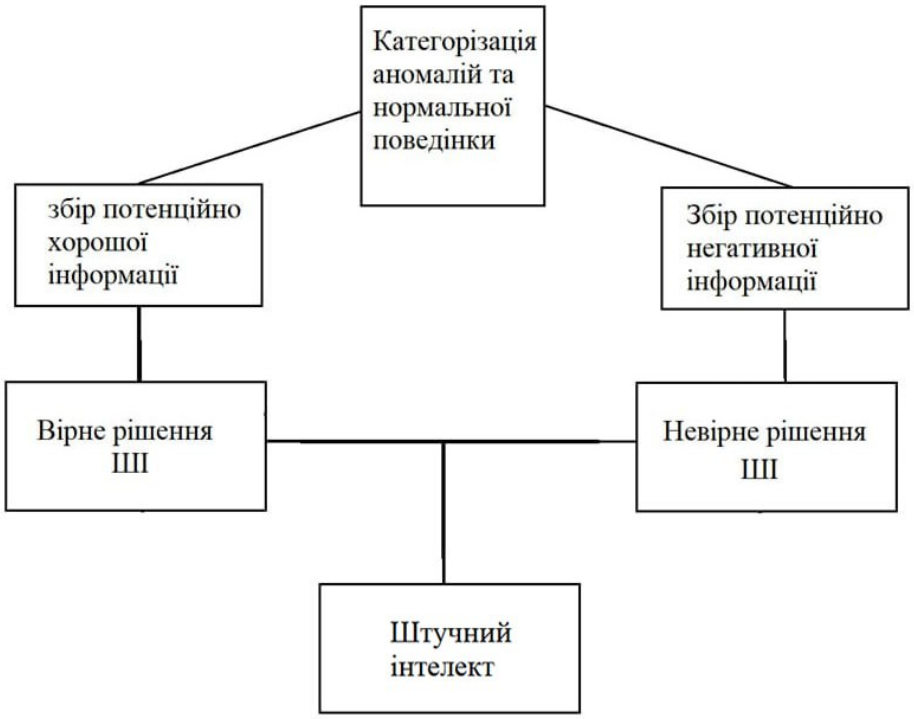

Рисунок 3.2 - Приклад схеми самостійного навчання найпростішого штучного інтелекту

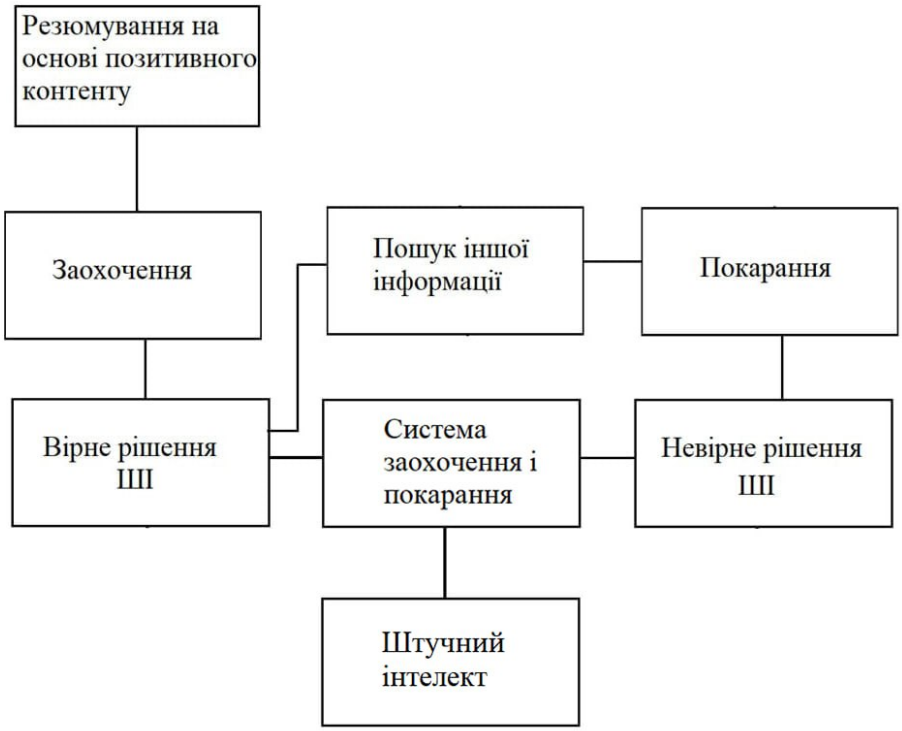

Рисунок 3.3 - Приклад схеми самостійного навчання найпростішого штучного інтелекту з допомогою додаткової інформації

Кожна система, яка тренується на виявлення загроз складається з трьох основних функціональних частин. Основні частини це:

- Розпізнавання образів
- Виявлення аномалій
- Керування ресурсами

Розпізнавання образів відповідає за аналіз інформації та визначення стандартних шаблонів або образів, які вказують на наявність конкретних загроз або нормальної поведінки. Вона може використовувати методи машинного навчання для виявлення характерних ознак кіберзагроз, таких як підозрілі шаблони трафіку, підозрілі активності користувачів тощо.

Виявлення аномалій це частина системи, яка спрямована на виявлення аномальних патернів або поведінки, яка може вказувати на наявність загроз. Вона використовує методи аналізу відхилень від норми для виявлення потенційно підозрілих дій або ситуацій.

Керування ресурсами відповідає за оптимальне використання доступних обчислювальних ресурсів для ефективного функціонування системи. Воно може включати в себе алгоритми розподілу ресурсів, моніторингу навантаження та автоматизацію процесів масштабування для забезпечення оптимальної продуктивності системи в умовах зростаючих обсягів даних та завдань.

3.2. Аналіз великих обсягів даних для виявлення аномалій та вразливостей

Для виконання аналізу великого обсягу даних зазвичай використовують Метод Ізольованого лісу.

Моделі ізольованих лісів - це алгоритми машинного навчання, які використовуються для виявлення аномалій у даних. Цей алгоритм особливо корисний у сфері кібербезпеки, де потрібно виявляти аномалії, підозрілі або незвичні патерни програмним шляхом. (див. рис. 3.4)

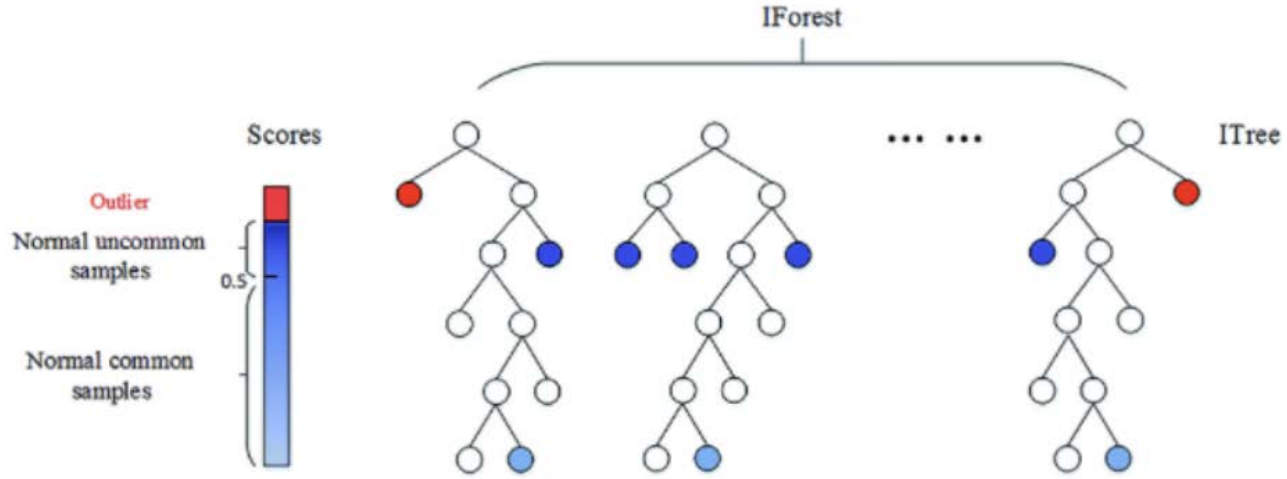

Рис. 3.4 – Схема моделі Ізольованого лісу

Основна ідея моделі ізольованого лісу полягає в тому, щоб ізолювати аномальні зразки даних від нормальних шляхом побудови лісу дерев випадкових рішень. Кожне дерево в лісі генерується шляхом вибору випадкової підмножини ознак і випадкового порогу для кожної ознаки. Після того, як дерево побудовано, алгоритм оцінює, які зразки з більшою ймовірністю потрапляють в лист. Аномальні екземпляри, які потрапляють в лист швидше, тобто потребують меншої кількості розщеплень, вважаються аномальними.

У сфері кібербезпеки модель ізольованого лісу може бути застосована для:

1. Виявлення вторгнень. Алгоритм може бути використаний для виявлення аномальної або підозрілої мережевої активності, яка може вказувати на потенційну атаку або вторгнення.

2. Виявлення шкідливого програмного забезпечення: Модель ізольованого лісу може аналізувати системну активність і виявляти незвичайні шаблони, які вказують на наявність шкідливого програмного забезпечення або вірусів.

3. Моніторингу веб-трафіку: За допомогою цього алгоритму можна аналізувати веб-трафік для виявлення незвичних або підозрілих дій, таких як спроби витоку даних або кібератаки.

4. Виявлення аномальної поведінки користувачів: Ця модель може бути використана для моніторингу поведінки користувачів і виявлення незвичайних

патернів, що вказують на порушення цілісності облікових записів або отримання несанкціонованого доступу.

Модель ізольованого лісу є потужним інструментом кібербезпеки, який допомагає швидко знайти підозрілу поведінку програм або аномалії під час їх виконання.

3.3 Використання нейронних мереж для прогнозування кібератак

Передбачення загроз за допомогою штучного інтелекту може оптимізувати виявлення загроз і створювати передові рішення для кібербезпеки, придатні для постійно мінливого ландшафту загроз. Ці системи штучного інтелекту можуть здійснювати самоконтроль і самонавчатися, застосовувати аналіз навіть у непередбачуваних ситуаціях і приймати рішення на основі власних спостережень.

Штучний інтелект може допомогти у виявленні та прогнозуванні загроз. Аналізувати великі масиви даних у режимі реального часу. Проактивний моніторинг мереж загроз вимагає курації великих обсягів структурованих і неструктурованих даних. Без штучного інтелекту цей процес вимагав би значної кількості часу і людських ресурсів. Штучний інтелект автоматизує моніторинг загроз, зменшує кількість людських помилок і полегшує їх виявлення.

Виявлення аномальних дій. Ідентифікація ризиків має вирішальне значення для прогнозування кібербезпеки за допомогою штучного інтелекту. Штучний інтелект може аналізувати зміни в мережі і використовувати закономірності для прогнозування. Поведінковий аналіз можна використовувати для виявлення аномальних дій всередині і зовні системи.

Прогнозування потенційних векторів атак. Методи та шляхи, які кіберзлочинці використовують для отримання доступу до системи або мережі називаються векторами атак. Штучний інтелект може відігравати важливу роль у прогнозуванні векторів атак на основі історичних даних. Методи машинного навчання можна використовувати для прогнозування майбутніх векторів атак, аналізуючи історичні дані та знаходячи закономірності й аномалії. Алгоритми

штучного інтелекту можуть аналізувати великі обсяги даних і виявляти приховані зв'язки між подіями, які неможливо виявити неозброєним оком. Чим більше даних ви надаєте, тим більше штучний інтелект навчається, а отже, з часом покращує свої можливості прогнозування.

3.4 Автоматизовані системи реагування на кіберзагрози

Існують різні автоматизовані системи реагування на кіберзагрози, які використовуються для виявлення, аналізу та реагування на потенційні атаки та інші кіберзагрози. Деякі приклади таких систем наведені нижче:

1. Системи управління інформацією та подіями безпеки (SIEM): SIEM-системи збирають, аналізують і візуалізують інформацію про події в режимі реального часу з різних джерел, включаючи журнали подій, системні журнали і мережевий трафік. Різні методи, такі як правила, евристика та машинне навчання, використовуються для виявлення аномальних моделей, підозрілих дій та потенційних загроз. (див. рис. 3.5)

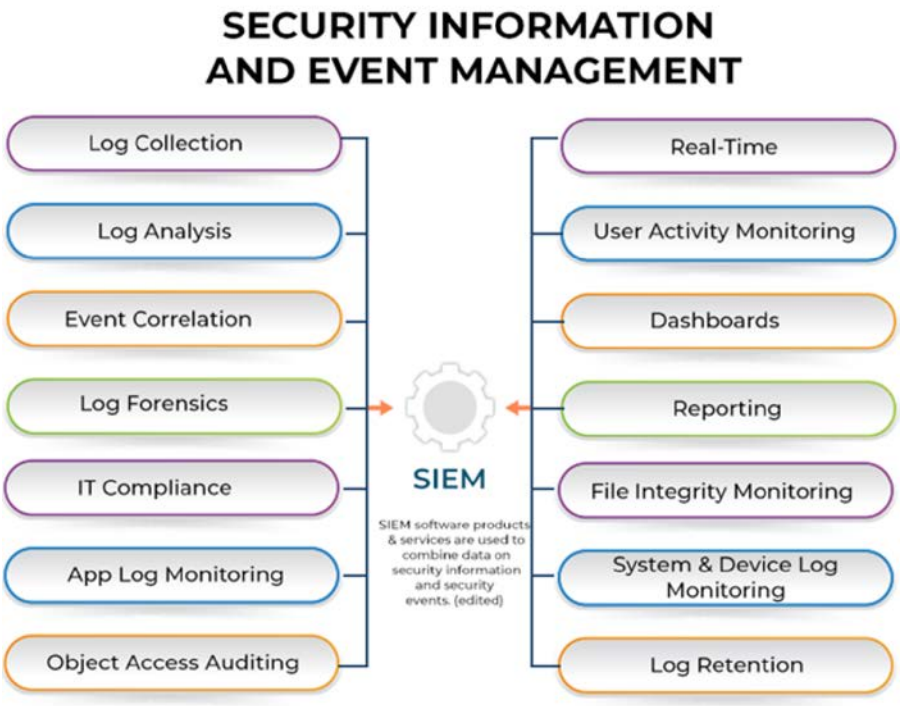

Рисунок 3.5 - Архітектура управління інформацією та подіями безпеки

На зображенні під номером 3.5 представлено компоненти та функції системи управління інформацією та подіями безпеки (SIEM). Нижче детально розписані основні компоненти та функції, зображені на малюнку.

– Збір та аналіз логів (Log Collection) - процес збирання даних із різних джерел, таких як мережеві пристрої, сервери та програми.

– Аналіз логів (Log Analysis) - аналіз зібраних даних для виявлення аномалій, підозрілих активностей або порушень політик безпеки.

– Кореляція подій (Event Correlation) - процес об'єднання та аналізу різних подій для виявлення складних загроз, які можуть бути неочевидними при окремому розгляді.

– Форензічний аналіз логів (Log Forensics) - поглиблений аналіз логів для відтворення подій та проведення розслідувань інцидентів безпеки.

– Відповідність IT-регламентам та аудит (IT Compliance) - переконання, що система відповідає внутрішнім та зовнішнім нормативним вимогам.

– Моніторинг логів застосунків (App Log Monitoring) - моніторинг логів програмного забезпечення для виявлення проблем безпеки та продуктивності.

– Аудит доступу до об'єктів (Object Access Auditing) - відстеження та аудит доступу до критичних об'єктів у системі для виявлення несанкціонованих доступів.

– Моніторинг у реальному часі (Real-Time) - здатність системи SIEM працювати в реальному часі для негайного виявлення та реагування на загрози.

– Моніторинг активності користувачів (User Activity Monitoring) - відстеження активності користувачів для виявлення потенційно небезпечних або несанкціонованих дій.

– Інформаційні панелі (Dashboards) - візуалізація даних та подій безпеки для швидкого прийняття рішень.

– Звітування (Reporting) - генерація звітів про стан безпеки та виявлені загрози.

– Моніторинг та збереження цілісності файлів (File Integrity Monitoring) - відстеження змін у важливих файлах для виявлення несанкціонованих модифікацій.

– Моніторинг логів систем та пристроїв (System & Device Log Monitoring) - постійний моніторинг логів від системних та мережевих пристроїв для забезпечення безпеки.

– Збереження логів (Log Retention) - зберігання логів протягом необхідного часу для проведення аналізу та відповідності нормативним вимогам.

SIEM використовує програмне забезпечення та послуги для об'єднання даних про події безпеки та їх аналізу. Ця система дозволяє виявляти нові загрози, аналізувати великі обсяги даних у реальному часі та знижувати ризики за допомогою інтеграції та кореляції інформації з різних джерел.

2. IDS/IPS (Intrusion Detection/ Prevention Systems): Ці системи виявляють аномальну або підозрілу мережеву активність і блокують атаки в режимі реального часу. Вони використовуються для виявлення мережевих вторгнень, шкідливого програмного забезпечення, використання вразливостей та інших загроз. (див. рис. 3.6)

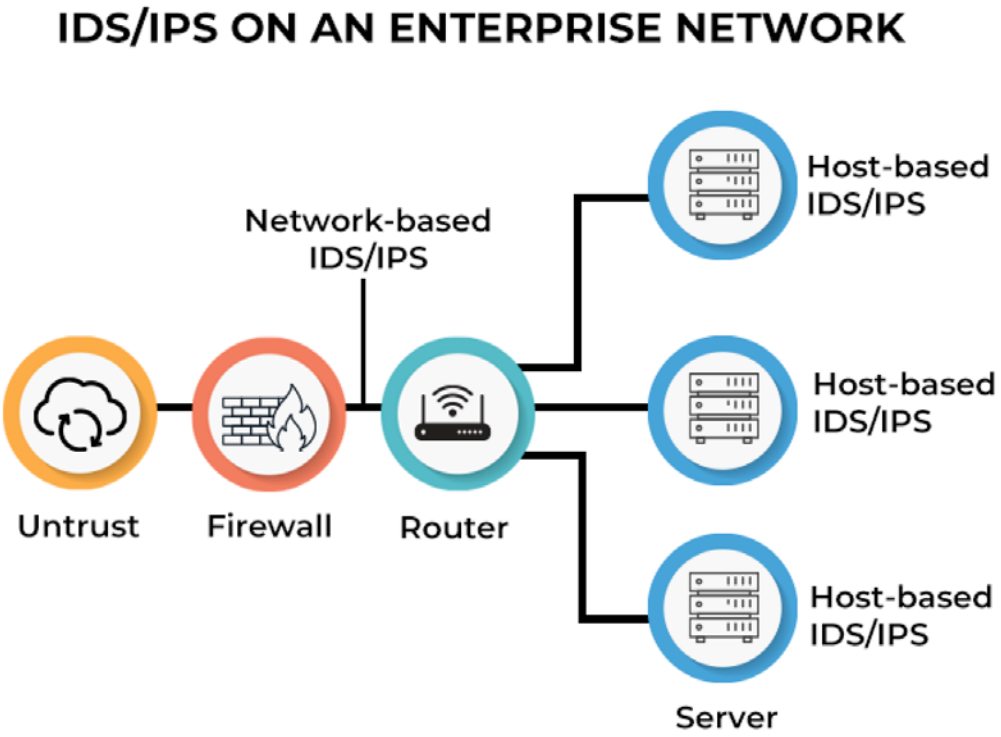

Рисунок 3.6 - IDS/IPS в корпоративній мережі

На зображенні під номером 3.6 представлена структура розгортання систем виявлення та запобігання вторгнень (IDS/IPS) в корпоративній мережі. Нижче детально розписані компоненти та їх взаємозв'язки:

– Ненадійна зона (Untrust) - це зовнішня мережа, яка несе потенційні загрози, зокрема Інтернет.

– Міжмережевий екран (Firewall) - захисний бар'єр між ненадійною зовнішньою мережею та внутрішньою корпоративною мережею. Він фільтрує трафік і блокує небажані з'єднання.

– Маршрутизатор (Router) - пристрій, що спрямовує трафік між різними сегментами мережі. У цьому контексті він також служить точкою інтеграції для мережевих систем IDS/IPS.

Типи систем IDS/IPS.

– Мережева система IDS/IPS (Network-based IDS/IPS) - розгортається в межах мережі для моніторингу та аналізу мережевого трафіку. Мережева IDS/IPS розташована між маршрутизатором і внутрішньою мережею.

– Хостова система IDS/IPS (Host-based IDS/IPS) - встановлюється безпосередньо на окремих пристроях, таких як сервери та робочі станції. Вона відстежує активність на конкретному хості та виявляє підозрілі дії на цьому пристрої.

– Мережева система IDS/IPS (Network-based IDS/IPS - розміщується після маршрутизатора, що дозволяє їй аналізувати весь вхідний та вихідний трафік з метою виявлення та запобігання атак на рівні мережі.

Системи IDS (Intrusion Detection System) та IPS (Intrusion Prevention System) відіграють ключову роль у забезпеченні кібербезпеки підприємств. Вони дозволяють виявляти та запобігати вторгненням шляхом моніторингу мережевого трафіку та активності на окремих пристроях.

Мережева IDS/IPS забезпечує аналіз і захист усього трафіку, що проходить через мережу, тоді як хостова IDS/IPS забезпечує більш детальний моніторинг і захист конкретних пристроїв.

Завдяки поєднанню мережевих та хостових систем IDS/IPS компанії можуть забезпечити багаторівневий захист своїх мереж від широкого спектра загроз.

3. Автоматизовані системи реагування на інциденти: Ці системи виконують автоматизовані дії для виявлення та реагування на інциденти. Наприклад, вони можуть автоматично блокувати IP-адреси зловмисників, відправляти на карантин скомпрометовані системи або надсилати сповіщення про інциденти безпеки відповідному персоналу.

4. Платформи розвідки загроз: Ці платформи збирають та аналізують дані про загрози з різних джерел і надають контекст для виявлення потенційних загроз. Деякі з них також можуть обмінюватися даними про загрози та автоматизувати процес виявлення загроз і реагування на них.

5. Системи виявлення та реагування на кінцеві точки (EDR): Системи EDR виявляють і реагують на загрози на рівні кінцевих точок, включаючи робочі станції, сервери і мобільні пристрої. Вони надають інформацію про активність, поширення та вплив загроз на окремі кінцеві точки. (див. Рис. 3.7)

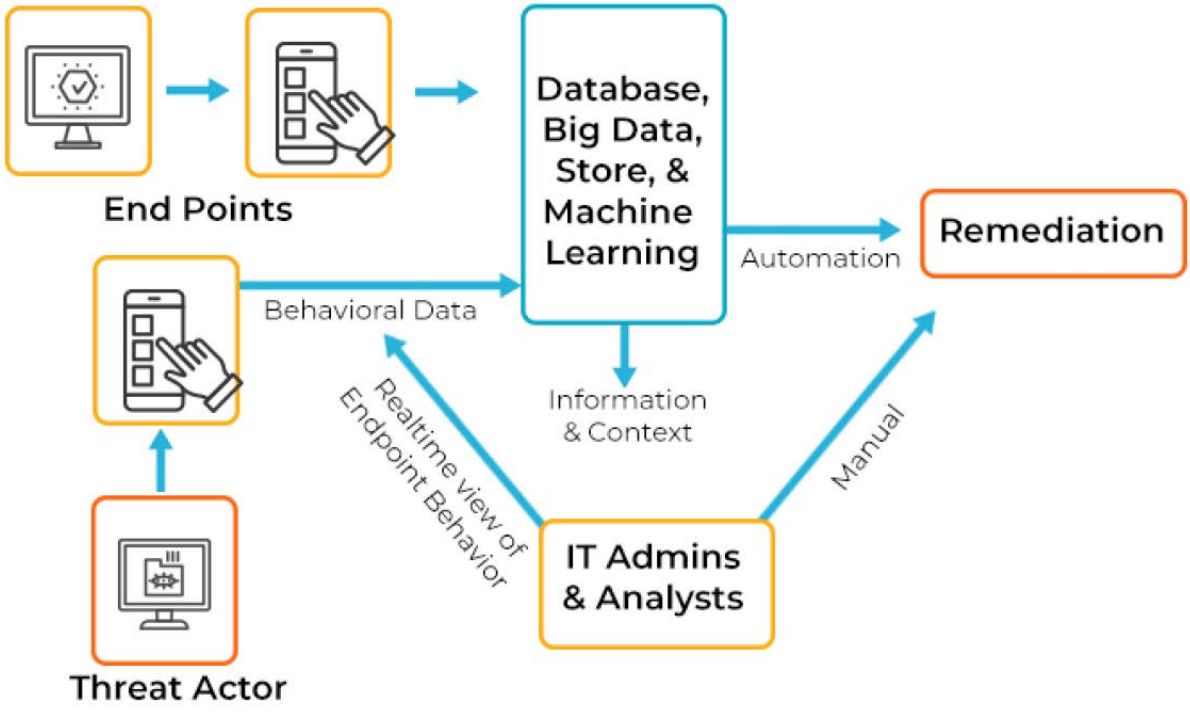

Рисунок 3.7 - Принцип роботи виявлення та реагування на кінцеві точки

Детальний опис компонентів та їх взаємодії з рисунка 3.7 наведено нижче.

– Кінцеві точки (End Points) - це пристрої, які підключені до мережі, такі як комп'ютери, мобільні телефони та інші пристрої користувачів. Вони є основними об'єктами моніторингу EDR-системи.

– Зловмисник (Threat Actor) - це суб'єкт або група, яка здійснює шкідливу діяльність, спрямовану на компрометацію кінцевих точок.

– Поведінкові дані (Behavioral Data) - дані, що збираються з кінцевих точок та відображають їхню поведінку. Ці дані використовуються для аналізу та виявлення підозрілих активностей.

– База даних, Великі дані, Сховище та Машинне навчання (Database, Big Data, Store, & Machine Learning) - центральна система, яка зберігає зібрані дані, аналізує їх за допомогою алгоритмів машинного навчання та виводить аналітичну інформацію. Ця система забезпечує автоматизацію процесів аналізу та виявлення загроз.

– Адміністратори IT та аналітики (IT Admins & Analysts) - група спеціалістів, яка отримує інформацію та контекст з центральної системи для подальшого аналізу та реагування. Вони можуть вручну втручатися в процес для детальнішого розслідування та вирішення інцидентів.

– Усунення наслідків (Remediation) - процес усунення виявлених загроз. Це може здійснюватися як автоматично (на основі аналізу даних), так і вручну (аналітиками та адміністраторами).

Поведенкові дані (Behavioral Data) передаються від кінцевих точок до центральної системи для аналізу.

Поточний перегляд поведінки кінцевих точок (Realtime View of Endpoint Behavior) забезпечує IT адміністраторам та аналітикам постійний доступ до даних про активність кінцевих точок.

Інформація та контекст (Information & Context) передаються від центральної системи до IT адміністраторів та аналітиків для подальшого аналізу.

Автоматизація (Automation) забезпечує автоматичне усунення загроз на основі результатів аналізу.

Ручне втручання (Manual) відбувається, коли IT адміністратори та аналітики втручаються для детальнішого розслідування та реагування на інциденти.

Система виявлення та реагування на кінцевих точках (EDR) є ключовим компонентом сучасної кібербезпеки, яка забезпечує моніторинг, аналіз та реагування на загрози на кінцевих точках мережі.

Зловмисники можуть атакувати кінцеві точки з метою отримання несанкціонованого доступу або заподіяння іншої шкоди.

EDR-системи збирають поведінкові дані з кінцевих точок і передають їх до центральної бази даних, де ці дані аналізуються за допомогою технологій машинного навчання.

IT адміністратори та аналітики використовують ці дані для отримання повного уявлення про активність кінцевих точок у реальному часі та реагують на виявлені загрози.

Процес усунення наслідків може бути як автоматизованим, так і здійснюваним вручну, залежно від складності загрози та необхідності детального розслідування.

Ці системи можуть допомогти підвищити швидкість та ефективність реагування на кіберзагрози за рахунок скорочення часу реагування та мінімізації впливу загроз на системи безпеки організації.

ВИСНОВКИ

Результатом виконання кваліфікаційної роботи є довершений аналіз можливих та існуючих застосувань штучного інтелекту у кібербезпеці. Робота проводить детальний аналіз застосування штучного інтелекту у кібербезпеці, висвітлюючи його переваги і недоліки. Вона враховує потенційні загрози і ризики, пов'язані з використанням нейромереж, машинного навчання та інших методів штучного інтелекту. Загалом у ній було розглянуто переваги та недоліки нейромереж, машинного навчання, його типи та потенційні загрози, що залежать від штучного інтелекту.

При написанні роботи було використано новітні дослідження та наукові статті, що доповідають про розвиток і практичне застосування сучасного штучного інтелекту.

Отримані теоретичні та практичні знання дозволяють зрозуміти сучасний стан використання штучного інтелекту в цій області. Вони можуть бути корисними для розробки стратегій кібербезпеки та впровадження заходів для захисту від кіберзагроз.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Parnashree Saha. The Role of Artificial Intelligence (AI) in Modern Cybersecurity. Encryption Consulting. (2024) URL: https://www.encryptionconsulting.com/the-role-of-artificial-intelligence-ai-in-modern-cybersecurity/

2. 上海市人工智能与社会发展研究会 文 | 绿盟科技集团股份有限公司 顾杜娟 杨鑫宜 王星凯 刘文懋 叶晓虎. 专题·人工智能安全|浅析人工智能技术在网络安全领域中的应用. SAASD 上海市人工智能与社会发展研究会. (2023) URL: https://saasd.org.cn/2023/07/06/专题·人工智能安全-浅析人工智能技术在网络安全/

3. Ящик О. Б., Симонов В. В., Іваненко Р. О. Забезпечення кібербезпеки в еру штучного інтелекту: аналіз технологічних підходів та стратегій для захисту інформації. Бізнес-інформ (2024). № 1_2024. С.81-86. URL: https://www.business-inform.net/export_pdf/business-inform-2024-1_0-pages-81_86.pdf

4. M. I. Jordan and T. M. Mitchell. Machine learning: Trends, perspectives, and prospects. Science: Special section – Artificial Intelligence (2015). 17 July 2015, Vol. 349, Issue 6245. Pp. 255-260. URL: https://www.cs.cmu.edu/~tom/pubs/Science-ML-2015.pdf

5. Ahmad F. Al Musawi. Introduction to Machine Learning (2018) URL: https://www.researchgate.net/publication/323108787_Introduction_to_Machine_Learning

6. Md. Zahangir Alom. Network Intrusion Detection for Cyber Security on Neuromorphic Computing System. International Joint International Joint Conference on Neural Networks (2017) 10.1109/IJCNN.2017.7966339 URL: https://www.researchgate.net/publication/315717823_Network_Intrusion_Detection_for_Cyber_Security_on_Neuromorphic_Computing_System

7. Ricardo Calderon. The Benefits of Artificial Intelligence in Cybersecurity. Economic Crime Forensics Capstones. (2019) 36. URL: https://digitalcommons.lasalle.edu/ecf_capstones/36/

8. Naveen Vemuri, Naresh Thaneeru, Venkata Manoj Tatikonda. Securing Trust: Ethical Considerations in AI for Cybersecurity. Journal of Knowledge Learning and Science Technology (2023) ISSN: 2959-6386 (Online), Vol. 2, Issue 2 Pp. 167-175. URL: https://jklst.org/index.php/home/article/view/142

9. Sarker, I. H.; Furhad, M. H.; Nowrozy, R. AI-Driven Cybersecurity: An Overview, Security Intelligence Modeling and Research Directions. SN Computer Science, (2021) URL: https://www.preprints.org/manuscript/202101.0457/v1

10. Wikipedia. History of artificial intelligence (2024) URL: https://en.wikipedia.org/wiki/History_of_artificial_intelligence

11. Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, Harsha Nori, Hamid Palangi, Marco Tulio Ribeiro, Yi Zhang. Sparks of Artificial General Intelligence: Early experiments with GPT-4. Microsoft Research. (2023) URL: https://arxiv.org/abs/2303.12712

12. Sergio Santoyo. A Brief Overview of Outlier Detection Techniques. Towards Data Science. (2017) URL: https://towardsdatascience.com/a-brief-overview-of-outlier-detection-techniques-1e0b2c19e561

13. Max Roser. The brief history of artificial intelligence: the world has changed fast — what might be next? Our World in Data. (2022) URL: https://ourworldindata.org/brief-history-of-ai

14. Chiradeep BasuMallick. What Is Endpoint Detection and Response? Spiceworks. (2022) URL: https://www.spiceworks.com/it-security/endpoint-security/articles/what-is-edr/

15. Hossein Ashtari. What Is an Intrusion Detection System (IDS)? Spiceworks. (2022) URL: https://www.spiceworks.com/it-security/network-security/articles/ids-vs-ips/

16. Remya Mohanan. Understanding the Architecture of Security Information and Event Management (SIEM) Spiceworks. (2022) URL: https://www.spiceworks.com/it-security/vulnerability-management/articles/what-is-siem/