УДК 004.43:004.855
DOI https://doi.org/10.32782/2663-5941/2023.5/38

Стаценко Д.В.
Київський національний університет технологій та дизайну

Стаценко В.В.
Київський національний університет технологій та дизайну

Злотенко Б.М.
Київський національний університет технологій та дизайну

Демішонкова С.А.
Київський національний університет технологій та дизайну

ДОСЛІДЖЕННЯ ПРОГРАМ НА ОСНОВІ ШТУЧНОГО ІНТЕЛЕКТУ В ЯКОСТІ КОМП'ЮТЕРНИХ ЗАСОБІВ ЗАХИСТУ ІНФОРМАЦІЇ

У роботі проведено дослідження застосування штучного інтелекту та машинного навчання, для удосконалення комп'ютерних систем аналізу та захисту від кіберзагроз. Зауважено, що за останні роки відбувається удосконалення та розробка нових типів шкідливого програмного забезпечення, в зв'язку з чим продовження використання традиційних систем захисту сьогодні – неефективно. Аналіз літератури показав, що глобальна вартість відновлення після типових порушень даних становить понад 4,45 мільйонів доларів США. Це вказує на необхідність додаткового удосконалення систем захисту із залученням технологій штучного інтелекту, з метою зниження додаткових витрати часу та ресурсів.

У статті запропоновано використання програм на основі штучного інтелекту для визначення та виявлення загроз з використанням машинного навчання для адаптації та передбачення появи нового або модифікації існуючого шкідливого програмного забезпечення.

Зауважено, що використання машинного навчання підвищує можливості автоматизації систем кібербезпеки та точності виявлення загроз. Штучний інтелект може збільшити кількість можливостей по визначенню загроз та попередження фахівців системи захисту про виявлення та блокування потенційного шкідливого програмного забезпечення.

В результаті проведеного досліджено представлено, що визначення загроз на основі штучного інтелекту підвищує ефективність та якість програм комп'ютерних систем захисту. Наведено моделі використання штучного інтелекту, які вказують на те, що покращується технології виявлення фішингу та шкідливих програм. Продемонстровано, що алгоритми машинного навчання можуть призвести до підвищення результативності визначення закономірностей та інших аномалій, що вказує на потенційні кіберзагрози. Проведення аналізу закономірностей та аномалій в даних, моделі штучного інтелекту можуть виявити й позначити потенційне шкідливе програмне забезпечення, додаючи нові можливості для систем кібербезпеки.

Ключові слова: *штучний інтелект, машинне навчання, кібербезпека, захист інформації, шкідливе програмне забезпечення.*

Постановка проблеми. Сьогодні все більше повсякденних процесів потребує використання нових цифрових систем, однак при цьому складність і удосконаленість кіберзагроз, з якими стикається людство, зростає. Традиційні заходи захисту цифрової безпеки стикаються з дедалі більшими труднощами та потребують постійного оновлення, щоб не відставати від цих загроз. Це призводить до необхідності в додаткових дослідженнях пов'язаних з більш досконалими і адаптивними рішеннями, такими як штучний інтелект (ШІ) і машинне навчання (МН) [1].

Штучний інтелект має можливість вчитися на основі отриманих даних і при цьому робити прогнози, що може позитивно вплинути на покращення кіберзахисту. МН, як підмножина ШІ, продемонструвало значний потенціал у покращенні виявлення кіберзагроз і реагування на них. Однак, незважаючи на це, застосування МН в кібербезпеці відноситься до нової сфери, яка потребує подальшого вивчення та розуміння [2].

Представлена робота спрямоване на дослідження штучного інтелекту та машинного

навчання в кібербезпеці з оглядом методів, які використовуються для аналізу загроз.

Аналіз останніх досліджень і публікацій. Кількість досліджень, пов'язаних із застосування штучного інтелекту в кібербезпеці, за останні роки збільшилось. Комплексні дослідження, щодо поточних застосувань штучного інтелекту в кібербезпеці, показали детальну статистику та розподіл робот пов'язаних з цим, також проводяться обговорення майбутніх напрямків досліджень. Дослідження надають детальне уявлення про кібербезпеку, керовану ШІ, з точки зору принципів і методів моделювання для інтелектуальних і автоматизованих служб пов'язаних з кібербезпекою. Незважаючи на перспективні розробки ШІ в кібербезпеці, потреба в подальших дослідженнях доцільна [3].

Крім того, розвиток Інтернету речей (IoT) розширив зону атаки для кіберзлочинців, що зумовило необхідність більш надійних і комплексних рішень безпеки. Програми кібербезпеки зараз розробляються з урахуванням безпеки Інтернету речей, забезпечуючи захист широкого кола пристроїв і мереж. Незважаючи на ці досягнення, сфера кібербезпеки продовжує стикатися зі значними проблемами. Зростаюча складність і масштаб кіберзагроз вимагають постійних інновацій і розвитку програм кібербезпеки. Штучний інтелект може значно покращити автоматичне реагування на інциденти, виявлення фішингу та виявлення шкідливих програм у сфері кібербезпеки [4].

У контексті освіти, систематичний огляд літератури масових відкритих онлайн-курсів (МВОК) з кібербезпеки показав, що існує потреба викладати ШІ та кібербезпеку разом. За даними IBM, з 2020 по 2023 рік вартість витоку даних зросла на 13,25% з 3,86 мільйона доларів до 4,45 мільйона доларів [5, 6]. Це вказує на необхідність додаткового удосконалення у ШІ, з метою зниження додаткових витрати часу та ресурсів.

Незважаючи на те, що штучний інтелект і машинне навчання демонструють потенціал у покращенні виявлення загроз і реагування на них, кількість досліджень присвячених цьому аспекту не є достатньою.

Метою статті є аналіз програм на основі штучного інтелекту та машинного навчання для використання в кібербезпеці для захисту інформації, а також механізмів аналізу загроз, які при цьому використовуються.

Виклад основного матеріалу. Штучний інтелект (ШІ) – це галузь інформатики, яка спрямована на створення систем, здатних виконувати завдання, які зазвичай потребують використання моделі людського інтелекту [7]. До цих завдань відноситься навчання та адаптація нової інформації, розуміння людської мови, розпізнавання моделей поведінки і прийняття рішень. Загалом ШІ розділяється на типи: 1) вузький ШІ, призначений для виконання конкретного завдання; 2) загальний ШІ – може розуміти, вивчати та застосовувати знання в широкому діапазоні завдань.

Кібербезпека передбачає захист комп'ютерних систем і мереж від розголошення інформації, крадіжки або її зміни (шифрування), захисту апаратного, програмного забезпечення, електронних даних, а також від порушення або неправильного спрямування послуг. ШІ може автоматизувати процес виявлення загроз і реагування на них, збільшуючи швидкість і ефективність. Він може вчитися, використовуючи для цього інформацію про минулі інциденти, визначати закономірності та передбачати майбутні загрози.

ШІ обробляє великі масиви даних на високій швидкості, що дозволяє виявляти загрози та реагувати на них у режимі реального часу. Адаптування до нових загроз дозволяє підвищити ефективність, в порівнянні з традиційними методи, які засновані на використанні бази відомих сигнатур загроз. Вищенаведене, у свою чергу, потенційно зменшить навантаження на фахівців з кібербезпеки, автоматизуючи базові та стандартні завдання, дозволяючи зосередитися лише на складних питаннях.

Однак використання штучного інтелекту в кібербезпеці також має проблеми. Однією з проблем є ризик помилкових спрацьовувань, коли законні дії помилково визначаються та ідентифікуються, як загрози, що призведе до непотрібних дій програм захисту і збоїв. Іншою проблемою є ризик маніпуляції або атаки зловмисників на системи штучного інтелекту. Відповідно забезпечення безпеки та цілісності систем ШІ є важливою проблемою. Також, при використанні штучного інтелекту в кібербезпеці необхідно розглядати питання етики та конфіденційності особистих даних користувачів, обробка яких виконується на різних етапах роботи систем захисту ШІ.

Машинне навчання (МН) – це підмножина ШІ, яка надає системам можливість автоматично навчатися та вдосконалюватися на основі досвіду без явного програмування [4, 8]. Він зосереджений на розробці програмного забезпечення, яке може отримувати доступ до даних і використовувати їх для самостійного навчання. Процес навчання включає такі елементи: 1) спостереження; 2) дані;

3) прямий досвід; 4) інструкції. Ці елементи дозволяють шукати закономірності в даних і приймати кращі рішення в майбутньому на основі прикладів, які надаються фахівцями з кібербезпеки.

У сфері кібербезпеки МН, завдяки своїй здатності швидко аналізувати величезні обсяги даних і виявляти аномалії, використовується для автоматизації виявлення загроз і боротьби з ними без втручання людини, таким чином скорочуючи час реакції та потенційно зменшуючи вплив загрози. МН також можна використовувати для прогнозування майбутніх загроз на основі раніше отриманих даних. Цей механізм орієнтований на дані намагається кількісно оцінити кіберризики, просуває методи висновку для аналізу моделей поведінки, зосереджується на створенні сповіщень про реакцію безпеки та оптимізує операції з кібербезпеки.

Нижче наведено декілька прикладів, які демонструють успішне впровадження МН у кібербезпеку. Компанія з кібербезпеки CrowdStrike використовує МН для автоматичного виявлення загроз і реагування на них. Їхні алгоритми МН аналізують дані з мільйонів систем у всьому світі, щоб ідентифікувати та блокувати потенційні загрози в реальному часі. Інший приклад є промислове застосування МН для захисту від кіберзагроз [8].

На основі загального потоку таких реалізацій представлення багаторівневої інтегрованої структури для машинного навчання в інтелектуальних службах кібербезпеки може працювати відповідно до описаного нижче, що включає в себе різні етапи обробки, від вихідних даних подій безпеки до кінцевих служб.

Загалом, потік може починатися з необроблених даних подій безпеки зібраних з різних джерел. Наступний етап передбачає підготовку цих даних та попередньої обробки для обробки алгоритмами машинного навчання. Далі відбувається застосування методів машинного навчання до цих даних, де модель вчиться ідентифікувати шаблони та аномалії, які можуть вказувати на потенційні загрози. Модуль постобробки та вдосконалення, який спрощує отримані знання відповідно до конкретних вимог шляхом включення предметно-специфічних знань. Далі йде модуль аналізу актуальності та оновлення моделі безпеки, який підтримує актуальність моделі безпеки шляхом завантаження найновіших шаблонів безпеки, керованих даними.

Останній модуль планування реагування та прийняття рішень, який приймає рішення на основі отриманої інформації та вживає необхідних заходів для запобігання системі від кібератак.

Машинне навчання в кібербезпеці також використовується для класифікації. Ця технологія забезпечує прогнозування зловмисності певного зразка з оцінкою впевненості. Ефективність цих моделей оцінюється на основі двох критеріїв: точності і результату. Позитивне виявлення від класифікатора зловмисного програмного забезпечення вказує на те, що модель передбачає зловмисний зразок на основі ознак, пов'язаних із відомими зловмисними зразками.

Розглянемо моделі, навчені аналізувати шкідливі файли. Справжній позитивний результат означає, що модель правильно визначила файл як шкідливий. Справжній негатив означає, що модель правильно визначила файл як нешкідливий. Помилковий результат означає, що модель неправильно визначила нешкідливий файл як шкідливий. Помилковий негатив означає, що модель неправильно визначила шкідливий файл як нешкідливий. Хоча справжні спрацьовування мають вирішальне значення для виявлення загроз і реагування на них, помилкові спрацьовування також є важливим показником ефективності. Помилкові спрацьовування негативно впливають на роботу всієї системи захисту, оскільки вони вимагають витрачати час і ресурси на дослідження кожного виявлення та можуть порушити роботу критичних програм при запуску автоматичних програм відновлення.

Під час коригування моделі системи захисту даних потрібно збалансувати показники справжніх і хибних позитивних результатів. Зниження порогу для справжніх позитивних результатів може збільшити кількість помилкових позитивних результатів, що призведе до втрати продуктивності. Метою розробки високопродуктивних моделей машинного навчання є максимізація ефективності виявлення шляхом збільшення справжніх позитивних виявлень і зменшення помилкових спрацьовувань. Цей баланс може бути складним, оскільки класифікатори зловмисного програмного забезпечення часто мають справжні позитивні показники близько 99%, збалансовані проти помилкових позитивних показників значно нижче 1%.

Таблиця 1

Помилкові позитивні та помилкові негативні результати

Передбачені значення	**Справжні значення**	
	Позитивні	**Негативні**
Позитивні	Справжні позитивні	Помилкові позитивні
Негативні	Помилкові негативні	Справжні негативні

Наступним етапом є розвідка загроз – збір і аналіз інформації про потенційні або поточні атаки, які загрожують організації. Концепція аналізу загроз передбачає аналіз та інтерпретацію даних для виявлення загроз, пошуку прогнозних індикаторів і впровадження захисних заходів. Роль штучного інтелекту в розвідці загроз полягає в автоматизації процесу збору, зберігання та аналізу даних.

Розвідка загроз на основі ШІ використовує машинне навчання та інші методи штучного інтелекту для аналізу шаблонів і виявлення аномалій, які вказують на потенційні загрози.

Для виявлення загроз сьогодні використовуються наступні бібліотеки штучного інтелекту та методів кодування. Scikit-learn, TensorFlow та PyTorch пропонують готові функції та інструменти для створення моделей машинного навчання. Ці бібліотеки можна використовувати для реалізації різних алгоритмів машинного навчання для виявлення загроз, наприклад випадковий ліс, дерево рішень та нейронні мережі.

Практики кодування також відіграють важливу роль у впровадженні аналізу загроз на основі ШІ. Перегляд коду, модульне тестування та постійна інтеграція, збільшує якість і надійність моделей ШІ.

Сьогодні використовується декілька наборів даних у сфері кібербезпеки, таких як NSL-KDD, DARPA, CAIDA, MAWI, ADFA IDS, CERT, EnronSpam, SpamAssassin, Malware Genome project, Virus Share, VirusTotal, Comodo, Contagio, Microsoft тощо. Дані набори даних містять приклади різних типів кібератак і можуть використовуватися для навчання та тестування продуктивності систем аналізу загроз з ШІ.

Наприклад, для завантаження набору даних NSL-KDD у фрейм для аналізу даних і машинного навчання за допомогою програмної бібліотеки pandas, використовується наступний код:

import pandas as pd

df = pd.read_csv('NSL-KDD/KDDTrain+.txt')

Впровадження методів ШІ та МН може значно підвищити продуктивність систем виявлення загроз. Наприклад, системи виявлення вторгнень можуть ідентифікувати різноманітні кіберзагрози та атаки, навіть невідомі атаки нульового дня, і реагувати в режимі реального часу на основі вимог користувачів.

Деякі додаткові бібліотеки AI, які можна використовувати для виявлення загроз:

– Keras: високорівневий API для нейронних мереж, написаний на Python і здатний працювати поверх TensorFlow, CNTK або Theano.

– Pandas: програмна бібліотека, яка пропонує структури даних і операції для роботи з числовими таблицями та часовими рядами, написана на мові програмування Python для обробки та аналізу даних.

– Numpy: бібліотека для мови програмування Python, що додає підтримку великих багатовимірних масивів і матриць разом із великою колекцією математичних функцій високого рівня для роботи з цими масивами.

Нижче наведено приклад програми реалізації бібліотеки для реалізації простої нейронної мережі та виявлення загроз:

```
import pandas as pd
import numpy as np
from keras.models import Sequentiel
from keras.layers import Dense
data = pd.read_csv('NSL-KDD/KDDTrain+.txt')
X = data.drop('label', axis=1)
Y=data['label']
model = Sequentiel()
model.add(Dense(12, input_dim=8, activation = 'relu'))
model.add(Dense(8, activation = 'relu'))
model.add(Dense(1, activation = 'sigmoid'))
model.compile(loss='binary_crossent', optimizer='adam', metrics=[accuracy])
model.fit(X, Y, epochs=150, bathc_size=10)
```

У представленій частині коду на першому етапі завантажується та попередньо оброблюється набір даних. На наступному етапі визначається проста нейронна мережу з одним вхідним шаром, одним прихованим шаром і одним вихідним шаром. Відбувається компіляція моделі за допомогою двійкової функції перехресних ентропійних втрат і оптимізатора *'Adam'*, а потім адаптація моделі до даних. Вихідні дані моделі визначаються, як передбачення того, чи є вхідні дані загрозою чи ні. В даному випадку наведено лише простий приклад фактичної реалізації, в залежності від конкретних вимог і контексту програма може бути більш комплексною і складною.

Автоматизоване реагування на інциденти (API) передбачає використання автоматизованих систем для виявлення інцидентів безпеки та реагування на них. Штучний інтелект може покращити API, забезпечивши виявлення загроз у реальному часі, автоматичне їх блокування і постійне навчання на основі минулих інцидентів. Нижче приведено приклад використання ШІ для вдосконалення API за допомогою Python і бібліотеки Scikit-learn для машинного навчання:

```
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import train_test_split
import pandas as pd
data = pd.read_csv('security_data.csv')
X = data.drop('threat_detected', axis=1)
Y = data['threat_detected']
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=42)
clf = RandomForestClassifier(n_estimators=100)
clf.fit(X_train, Y_train)
y_pred = clf.predict(X_test)
for i in range(len(Y_pred)):
    if Y_pred[i] == 1:
        print(f'Threat detected in data point {X_test.iloc[i]}. Initiating automated response.')
```

Коли система ШІ виявляє загрозу, відбувається запуск повідомлення про те, що загроза була визначена у відповідному пакеті даних і відбувається автоматична реакція на неї.

Також ШІ може покращити виявлення фішингу, аналізуючи електронні листи або веб-сайти та визначаючи характеристики, які можуть вказувати на спробу фішингу. Нижче наведено приклад використання ШІ для виявлення фішингу за допомогою Python і Natural Language Toolkit (NLTK) для аналізу тексту:

```
import nltk
from sklearn.featureextraction.text import  CountVectorizer
from sklearn.naive_bayes import MultinomialNB
from sklearn.model_selection import train_test_split
import pandas as pd
data pd.read_csv('phishing_data.csv')
X = data['email_text']
Y = data['is_phishing']
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=42)
Vectorizer = CountVectorizer()
X_train_counts = vectorizer.fit_transform(X_train)
clf = MultinomialNB()
clf.fit(X_train_counts, Y_train)
X_test_counts = vectorizer.transform(X_test)
Y_pred = clf.predict(X_test_counts)
for I in range(len(Y_pred)):
    if Y_pred[i] == 1:
        print(f'Phishing attempt detected in email: {X_test.iloc[i]}.')
```

Система ШІ виявляє спробу фішингу на основі текстового вмісту електронних листів. Коли виявляється спроба фішингу, система запускає сповіщення про те, в якому листі знаходиться шкідливе повідомлення.

Шкідливе програмне забезпечення включає віруси, хробаки, трояни, програми-вимагачі та інші шкідливі програми. Штучний інтелект може покращити виявлення шкідливого програмного забезпечення, аналізуючи характеристики програмного забезпечення та ідентифікуючи функції, які можуть вказувати на його присутність. Нижче наведено приклад використання ШІ для покращення виявлення шкідливого програмного забезпечення за допомогою Python і бібліотеки TensorFlow для глибокого навчання:

```
import tensorflow as tf
from sklearn.model_selection train_test_split
import pandas as pd
data pd.read_csv('malware_data.csv')
X = data.drop('is_malware', axis=1)
Y = data['is_malware']
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=42)
model = tf.keras.models.Sequential([
    tf.keras.layers.Dense(128, activation 'relu'),
    tf.keras.layers.Dense(64, activation 'relu'),
    tf.keras.layers.Dense(1, activation 'sigmoid')
])
model.compile(optimizer='adam', loss='binary_crossent', metrics=['accuracy'])
model.fit(X_train, Y_train, epochs=5)
y_pred model.predict(X_test)

for i in range(len(y_pred)):
    if y_pred[i] >= 0.5:
        print(f'Malware detected in software: {X_test.iloc[i]}.')
```

Система ШІ виявляє шкідливі програми на основі характеристик програмного забезпечення і при виявленні загрози вона запускає автоматичне сповіщення.

Висновки. Представлена робота демонструє значний потенціал використання штучного інтелекту для посилення заходів в галузі кібербезпеки. Аналіз загроз на основі ШІ може покращити виявлення загроз і час реагування, забезпечуючи надійний захист від кібератак. Крім того, застосування штучного інтелекту в кібербезпеці не тільки призводить до збільшення випадків виявлення загроз, але й має позитивний вплив автоматизованого реагування на інциденти. Наведенні моделі можуть аналізувати загрози та реагувати на них у реальному часі, скорочуючи час, необхідний для зниження ефективності потенційних кібератак.

Визначено що ШІ також може покращити виявлення фішингу та шкідливих програм. Аналізуючи закономірності та аномалії в даних, моделі штучного інтелекту можуть виявити й позначити потенційне шкідливе програмне забезпечення, додаючи нові можливості для систем кібербезпеки.

Список літератури:

1. Jordan M. I., Mitchell T. M. Machine learning: Trends, perspectives, and prospects. *Science*, vol. 349, no. 6245, pp. 255-260, 2015.

2. LeCun Y., Bengio Y., Hinton G. Deep learning. *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.

3. Mehedi Hasan M. A., Salah K., Jayaraman R., Iqbal Hossain M., Alhamad M., Guizani A. Cybersecurity data science: an overview and future direction. *Journal of Big Data*, vol. 7, no. 1, pp. 1-25, 2020. URL: https://journalofbigdata.springeropen.com/counter/pdf/10.1186/s40537-020-00318-5.pdf.

4. Sarker I. H. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, vol. 2, no. 2, pp. 1-25, 2021. URL: https://dx.doi.org/10.1007/s42979-021-00592-x.

5. Torska N. 500+ Data Breach Statistics: Hystory, Cost & Preventing. *MarketSplash.* 2023. URL: https://marketsplash.com/data-breach-statistics/.

6. Cost of a Data Breach Report 2023. *IBM*. 2023. URL: https://www.ibm.com/reports/data-breach

7. Zeadally S., Adi E., Baig Z., Khan I. A. Harnessing Artificial Intelligence Capabilities to Improve Cybersecurity, *IEEE Access*, vol. 8, pp. 31830-31850, 2020. URL: https://dx.doi.org/10.1109/ACCESS.2020.2968045.

8. Sarker I. H., Furhad M. H., Nowrozy R. AI-Driven Cybersecurity: An Overview, Security Intelligence Modeling and Research Directions, *SN Computer Science*, vol. 2, no. 2, pp. 1-25, 2021. URL: https://dx.doi.org/10.1007/s42979-021-00557-0.

Statsenko D.V., Statsenko V.V., Zlotenko B.M., Demishonkova S.A. RESEARCH OF PROGRAMS BASED ON ARTIFICIAL INTELLIGENCE AS COMPUTER TOOLS FOR INFORMATION PROTECTION

The paper examines the use of artificial intelligence and machine learning to improve computer systems for analyzing and protecting against cyber threats. It has been noted that in recent years there has been improvement

and development of new types of malicious software, which is why continuing to use traditional protection systems today is ineffective. A literature review found that the global recovery cost of a typical data breach is over US $4.45 million. This indicates the need for additional improvement of protection systems with the involvement of artificial intelligence technologies, with the aim of reducing additional consumption of time and resources.

The article proposes the use of programs based on artificial intelligence to identify and detect threats using machine learning to adapt and predict the appearance of new or modification of existing malicious software.

It is noted that the use of machine learning increases the possibilities of automation of cyber security systems and the accuracy of threat detection. Artificial intelligence can increase the number of opportunities to identify threats and alert security professionals to detect and block potential malware.

As a result of the research, it is presented that the identification of threats based on artificial intelligence increases the efficiency and quality of programs of computer protection systems. Artificial intelligence usage patterns are presented, indicating that phishing and malware detection technologies are improving. It has been demonstrated that machine learning algorithms can lead to an increase in the effectiveness of identifying patterns and other anomalies that indicate potential cyber threats. By analyzing patterns and anomalies in data, AI models can detect and flag potential malware, adding new capabilities to cybersecurity systems.

Key words: *artificial intelligence, machine learning, cyber security, information protection, malicious software.*