DOI: 10.58254/viti.6.2024.08.105

УДК 004.738.5:355/359 (021)

Мальцева І. Р. ORCID: 0000-0001-6073-4637 (ВІТІ ім. Героїв Крут)
Черниш Ю. О. ORCID: 0000-0002-6626-5656 (ВІТІ ім. Героїв Крут)
Процюк Ю. О. ORCID: 0000-0001-5193-3669 (ВІТІ ім. Героїв Крут)

АНАЛІЗ АЛГОРИТМІВ РАННЬОГО ВИЯВЛЕННЯ КІБЕРАТАК НА МЕРЕЖІ З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ

Захист критичної інфраструктури та національна безпека посилюються завдяки безпеці та надійності мереж. В цих мережах циркулює різна інформація з різними грифами від відкритої до закритої. Наслідки кібератак на ці мережі можуть, бути серйозними, включаючи шкоду репутації, фінансові втрати, операційні перебої та витік даних. Традиційні методи безпеки, такі як брандмауери та антивірусне програмне забезпечення, стають все менш ефективними проти сучасних і постійно мінливих кіберзагроз. Як наслідок, потужні системи виявлення мережевих вторгнень (IDS) стали незамінними для проактивного виявлення та пом'якшення наслідків кібератак.

Машинне навчання стало життєздатним методом для створення адаптивних засобів виявлення вторгнень, які можуть виявляти нові та складні шаблони атак. Навчаючись на величезних маркованих наборах даних мережевого трафіку, моделі ML можуть розуміти тонкі закономірності і диференціальні ознаки нормальних і аномальних або зловмисних потоків трафіку. Це дозволяє виявляти можливі кіберзагрози та вторгнення, які традиційні IDS на основі сигнатур не можуть виявити. Виділення дискримінаційних ознак і навчання відповідних моделей класифікації з таких даних є непростим завданням.

У представленому дослідженні проведено аналіз ефективності алгоритмів ML для виявлення кібератак, зокрема розподілених атак на відмову в обслуговуванні DDoS, у даних мережевого трафіку. У представленому дослідженні система виявлення мережевих атак розроблена з використанням моделей ML і глибинного навчання (DL) та експериментується на наборі даних CICIDS2017.

Основними завданнями дослідження є розроблення стратегії вилучення цінної інформації з необроблених мережевих потоків; вивчення впливу підготовки даних на частоту хибних спрацьовувань; проведення порівняльного аналізу моделей ML для ідентифікації кібератак.

Основною метою дослідження є забезпечення розуміння розробки надійної адаптивної системи виявлення мережевих вторгнень з використанням підходів ML, які підвищують можливості кібербезпеки та захищають від майбутніх кібератак.

Ключові слова: *машинне навчання, виявлення кібератак, розподілена атака на відмову в обслуговуванні, виявлення мережевих вторгнень, нейронні мережі.*

I. Maltseva, Y. Chernish, Y. Protsiuk Development of algorithms for early detection of cyberattacks on networks using machine learning

Critical infrastructure protection and national security are enhanced by the security and reliability of networks. Various types of information circulate on these networks, ranging in classification from open to closed. The consequences of cyberattacks on these networks can be severe, including reputational damage, financial loss, operational disruption and data leakage. Traditional security methods, such as firewalls and anti-virus software, are becoming less effective against modern and ever-changing cyber threats. As a result, powerful network intrusion detection systems (IDS) have become indispensable for proactive detection and mitigation of cyber attacks.

Machine learning has become a viable method for creating adaptive intrusion detection tools that can detect new and complex attack patterns. By learning from huge labelled network traffic datasets, ML models can understand the subtle patterns and differentiating features of normal and abnormal or malicious traffic flows. This allows them to detect possible cyber threats and intrusions that traditional signature-based IDSs cannot detect. Extracting discriminative features and training appropriate classification models from such data is a challenging task.

In the presented study, we analyse the effectiveness of ML algorithms for detecting cyberattacks, in particular distributed denial of service (DDoS) attacks, in network traffic data. In the presented study, a network attack detection system is developed using ML and deep learning (DL) models and experimented on the CICIDS2017 dataset.

The main objectives of the study are to develop a strategy for extracting valuable information from raw network streams; to study the impact of data preparation on the false positive rate; and to conduct a comparative analysis of ML models for cyberattack detection.

The main goal of the study is to provide an understanding of the development of a reliable adaptive network intrusion detection system using ML approaches that increase cybersecurity capabilities and protect against future cyberattacks.

Keywords: *machine learning, cyberattack detection, distributed denial of service attack, network intrusion detection, neural networks.*

Постановка проблеми. Захист критичної інфраструктури та національна безпека посилюються завдяки безпеці та надійності мереж. В цих мережах циркулює різна інформація з різними грифами від відкритої до закритої. Надання послуг урядами, компаніями і приватними особами призвело до феноменального сплеску використання Інтернету. Незважаючи на очевидні переваги, цей взаємозв'язок робить мережі вразливими до кібернетичних наслідків і атак у широких масштабах [1]. Наслідки кібератак на ці мережі можуть бути серйозними, включаючи шкоду репутації, фінансові втрати, операційні перебої та витік даних. Традиційні методи безпеки, такі як брандмауери та антивірусне програмне забезпечення (АВПЗ), стають все менш ефективними проти сучасних і постійно мінливих кіберзагроз [2]. Як наслідок, потужні системи виявлення мережевих вторгнень (IDS) стали незамінними для проактивного виявлення та пом'якшення наслідків кібератак [3].

Машинне навчання (ML) стало життєздатним методом для створення адаптивних засобів виявлення вторгнень, які можуть виявляти нові та складні шаблони атак [4]. Навчаючись на величезних маркованих наборах даних мережевого трафіку, моделі ML можуть розуміти тонкі закономірності і диференціальні ознаки нормальних і аномальних або зловмисних потоків трафіку [5]. Це дозволяє виявляти можливі кіберзагрози та вторгнення, які традиційні IDS на основі сигнатур не можуть виявити. Однак застосування моделей ML в індустрії кібербезпеки є досить складним. Дані мережевого трафіку надзвичайно різноманітні та незбалансовані, і лише невеликий відсоток з них вказує на справжні інциденти атак [6]. Виділення дискримінаційних ознак і навчання відповідних моделей класифікації з таких даних є непростим завданням. Крім того, IDS повинна підтримувати надзвичайно низький рівень помилкових спрацьовувань, щоб не перевантажувати команди безпеки неправильними оповіщеннями [7].

Аналіз останніх досліджень і публікацій. За останні роки численні дослідники вивчали методи ML для виявлення вторгнень та аномалій у мережі. Автори публікації [8] проаналізували різні підходи ML, такі як Дерева рішень (DT), опорно-векторні машини (SVM), випадкові ліси (RF) та наївний Баєс (NB), на наборах даних NSL-KDD та KDDCup'99 для виявлення мережевих вторгнень.

Було виявлено, що модель RF досягла точності 99,9% на наборі даних NSL-KDD. В роботі [9] запропоновано новий підхід на основі кластеризації, який групує різні типи атак за даними мережевого трафіку в окремі кластери. Це дозволяє розробляти індивідуальні моделі виявлення вторгнень для кожної категорії атак.

Проаналізувавши різні типи мережевих атак, які повинні виявляти моделі ML, включаючи DDoS [10], SQL-ін'єкції [11], виявлення "безпечної оболонки" грубою силою (SSH) [12], атаки типу "людина посередині" та інші, автори [13] спеціально вивчали питання безпеки бездротових мереж і підкреслили важливість виявлення активних атак, таких як DDoS, у відкритих середовищах [14].

Набори даних (KDDCup1999 та NSL-KDD) були широко застосовані для оцінки мережевих IDS [15–20]. Однак, слід зазначити, що ці набори даних є досить застарілими і не зовсім точно відображають сучасні мережеві атаки.

Мета та завдання дослідження. Метою дослідження є вивчення ефективності алгоритмів ML для виявлення кібератак, зокрема розподілених атак на відмову в обслуговуванні DDoS, у даних мережевого трафіку. Проведення порівняльного аналізу моделей ML для ідентифікації кібератак.

Виклад основного матеріалу. Для експериментів використовується новий набір даних CICIDS2017, який є найсучаснішим, що включає найпоширеніші потоки атак. Досліджено та

оцінено кілька методів ML, включаючи DT, штучні нейронні мережі (ANN) та ансамблеві підходи (EL).

Для виявлення мережевих атак оцінюються класичні моделі ML, EL та підходи DL. Класичні підходи ML включають SVM, DT і метод K-найближчих сусідів (KNN), підходи EL включають RF, метод AdaBoost і XGBoost, а методи DL – архітектуру багатошарового перцептрона (MLP) і нову модель Deep MLP. Дві архітектури щільних шарів використовуються в Deep MLP з функціями активації ReLU та сигмоїдної активації. Представлене дослідження сприяло бінарній класифікації зразків мережевого трафіку як атакуючих, так і нормальних.

Для навчання та тестування моделей було використано набір даних CICIDS2017, створений Канадським інститутом кібербезпеки (CIC). CICIDS2017 містить велику базу даних мережевого трафіку для експериментів.

Особливістю CICIDS2017 є представлення широкого спектра сценаріїв атак – DoS і DDoS-атак, вебвторгнень, таких як SQL-ін'єкції та "міжсайтовий скриптинг" (XSS), інфільтраційні атаки та діяльність бот-мереж. Створений навмисно для відтворення реальних загроз, цей набір даних є ресурсом для дослідників, які розробляють та оцінюють системи виявлення вторгнень та інші рішення з кібербезпеки. Дана база даних розподілена на такі сегменти, як Normal traffic, Fuzzers, Reconnaissance, Analysis, Backdoors, DoS attacks, Exploits, Shellcode, Generic attacks та Worms.

Нехай набір даних (S_1, S_2, ..., Sn) складається з n вибірок, де кожна вибірка S_i представлена вектором ознак (f_1, f_2, ..., f_{47}, T, C). Ознаки від f_1 до f_{47} представляють різні характеристики мережевого трафіку, T – це мітка типу атаки, а C – бінарна мітка мережевого трафіку, яка вказує на те, чи є зразок нормальним, чи є він атакою. Мітка типу атаки T може приймати одне з значень: Analysis, Backdoors, DoS attacks, Fuzzers, Generic attacks, Shellcode, Reconnaissance, Exploits, Worms або Normal traffic. З вибірки S_j мітка мережевого трафіку представляє статус трафіку, атакованого або нормального. Основна мета моделей класифікації – віднести трафік до класів атакованого або звичайного трафіку. Більше того, для зразків, ідентифікованих як атаки, модель повинна визначати конкретний тип атаки, пов'язаний з "S_j". Комплексна архітектура для виявлення мережевих загроз представлена на рис. 1.

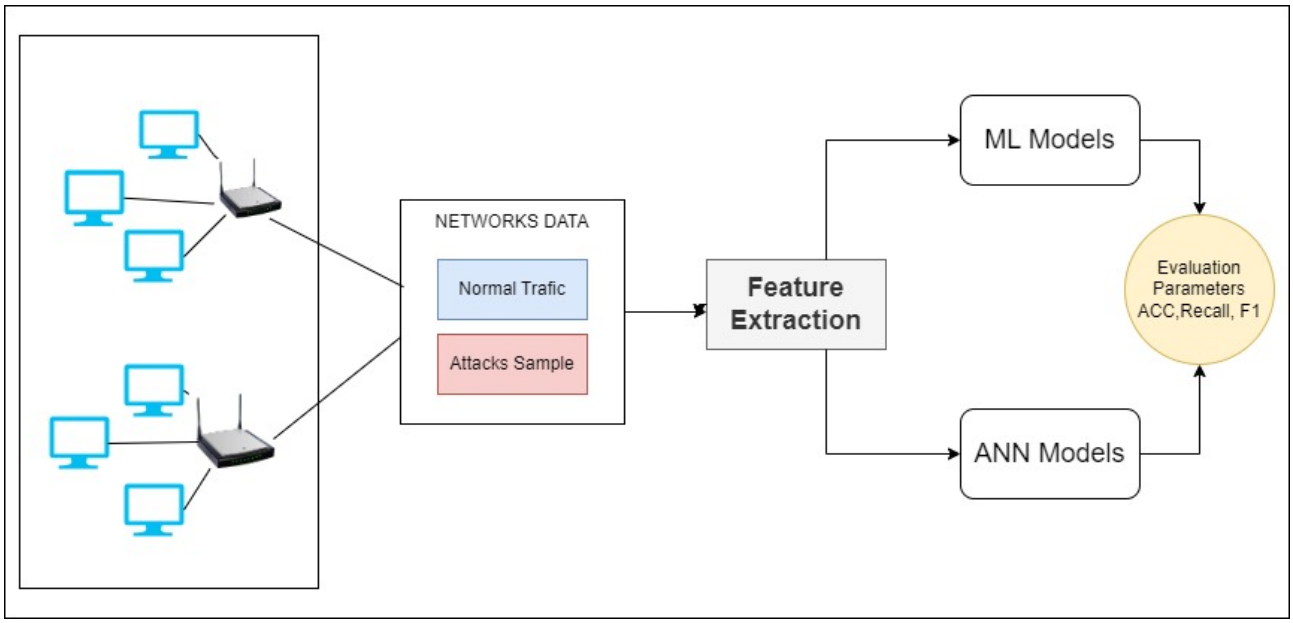

Рис. 1. Комплексна архітектура для виявлення мережевих загроз

Виявлення мережевих атак є критично важливим завданням у сфері кібербезпеки і вимагає спеціальних знань. ML пропонує потужний підхід для виявлення та пом'якшення Exploits і типових кібератак, використовуючи свою здатність аналізувати великі та складні набори даних для виявлення шаблонів і аномалій, які можуть вказувати на зловмисну діяльність. У контексті виявлення Exploits моделі ML можна навчити розпізнавати сигнатури відомих Exploits шляхом аналізу функцій, отриманих із мережевого трафіку, системних журналів і поведінки програми. Наприклад, алгоритми контрольованого навчання можна використовувати для класифікації шаблонів трафіку як доброякісних або шкідливих на основі позначених наборів даних, що дозволяє виявити відомі Exploits з високою точністю. Класичні алгоритми, такі як SVM, ANN та DT, досягли успіху в цій галузі. SVM вміють знаходити оптимальні гіперплощини, що робить їх ідеальними для великих наборів ознак. KNN посилається на сусідні зразки для прогнозування, в той час як DT надають швидкі, зрозумілі рішення шляхом розбиття даних на частини. З розвитком кіберзагроз з'являється простір для надійних методів виявлення. XGBoost, що використовує градієнтне підсилення та регуляризацію, вирізняється низькою чутливістю до перенавчання. MLP, що застосовується для нелінійного мапування, використовує зворотне поширення для вивчення складних патернів у прихованих шарах.

Нова архітектура Deep MLP з активацією ReLU та сигмоїдною функцією активації фіксує мінімальні зміни тенденції мережевого трафіку для точного виявлення атак. XGBoost, і Deep MLP пропонують надійні інструменти для ML та DL, що дозволяють автоматично розпізнавати складні закономірності у величезних масивах даних. Методи регуляризації гарантують, що моделі зберігають ефективність у різних сферах застосування, підвищуючи їхню стійкість і корисність у кібербезпеці.

ML особливо ефективний у виявленні типових кібератак, таких як phishing, DDoS-атаки та зараження шкідливим програмним забезпеченням. Використовуючи методи виявлення аномалій, ML може виявляти відхилення від нормальної поведінки, які можуть означати триваючу атаку. Моделі неконтрольованого навчання, такі як кластерізація та автокодери, можуть ідентифікувати нові атаки або атаки нульового дня, які не відповідають жодним відомим сигнатурам, позначаючи незвичайні шаблони або викиди в даних.

У реальних додатках комбінація методів машинного навчання часто використовується для створення надійних і адаптивних систем безпеки. Гібридні моделі, які об'єднують методи виявлення як на основі сигнатур, так і на основі аномалій, забезпечують комплексний механізм захисту, здатний протистояти широкому спектру загроз. Крім того, постійне навчання та адаптація мають вирішальне значення, оскільки моделі ML мають регулярно оновлюватися новими даними та сигнатурами атак, щоб залишатися ефективними проти кіберзагроз, що постійно змінюються. Цей підхід не тільки покращує здатність виявляти кібератаки та реагувати на них у режимі реального часу, але й знижує рівень помилкових спрацьовувань, тим самим покращуючи загальну безпеку інформації.

Показники оцінювання

Система виявлення мережевих атак оцінюється за допомогою оціночних метрик, що виводяться з матриці заплутаності. Матриця заплутаності, таблиця 2×2, обчислює істинні спрацьовування (правильно ідентифіковані атаки, правильно ідентифікований нормальний трафік), хибні спрацьовування (пропущені атаки, неправильно ідентифіковані атаки).

Прогностична здатність кожної моделі оцінюється за показником продуктивності (Accuracy) та розраховується як частка правильно класифікованих зразків від загальної кількості зразків. Відклик (Recall), також відомий як істинно позитивний показник (True Positive rate), відображає здатність моделі виявляти атаки, обчислюючи частку правильно ідентифікованих атак, а правильний прогноз представлено метрикою Precision, що вимірює надійність позитивних прогнозів, оцінюючи частки справжніх атак серед зразків, позначених

як атаки. Показник F1 дає збалансовану оцінку ефективності моделі, враховуючи як Precision, так і можливість запам'ятовування.

Дослідження та результати

Апаратна специфікація для розробки системи виявлення мережевих вторгнень включає процесор Intel Core i7 (11-го покоління), 16 ГБ оперативної пам'яті та жорсткий диск 1 ТБ, а також операційну систему Ubuntu 20.04.4. Мова програмування Python використовувалася для реалізації моделі ML з бібліотекою Keras 2.3.1 для підтримки TensorFlow 2.2.0 для розробки та навчання нейромережевих моделей.

Ефективність моделей ML

Результати роботи різних моделей ML у виявленні мережевих атак представлені в таблиці 1, 2 та на рис. 2, 3. Моделі оцінюються за допомогою запам'ятовування метрик, продуктивності (Accuracy), точності (Precision) та збалансованої оцінки ефективності моделі (F1). Відклик (Recall) відображає здатність моделі знаходити відповідні випадки, а правильний прогноз представлено метрикою Precision.

Accuracy на рівні 99,05 %, досягнута класифікатором DT, та високою збалансованою оцінкою ефективності моделі на рівні 99 %, що свідчить про ефективність моделі у виявленні мережевих атак. Модель SVM також продемонструвала високу Accuracy на рівні 95,17 та отримала збалансовану оцінку ефективності моделі на рівні 94 % відповідно.

Таблиця 1

Оцінка ефективності методів ML

Model	Recall	Precision	Accuracy	F1-measure
DT	99	99	99,05	99
SVM	93	96	95,17	94
RF	99	99	98,96	99
AdaBoost	97	98	97,87	98
XGBoost	97	98	98,08	98
MLP	96	98	97,47	97

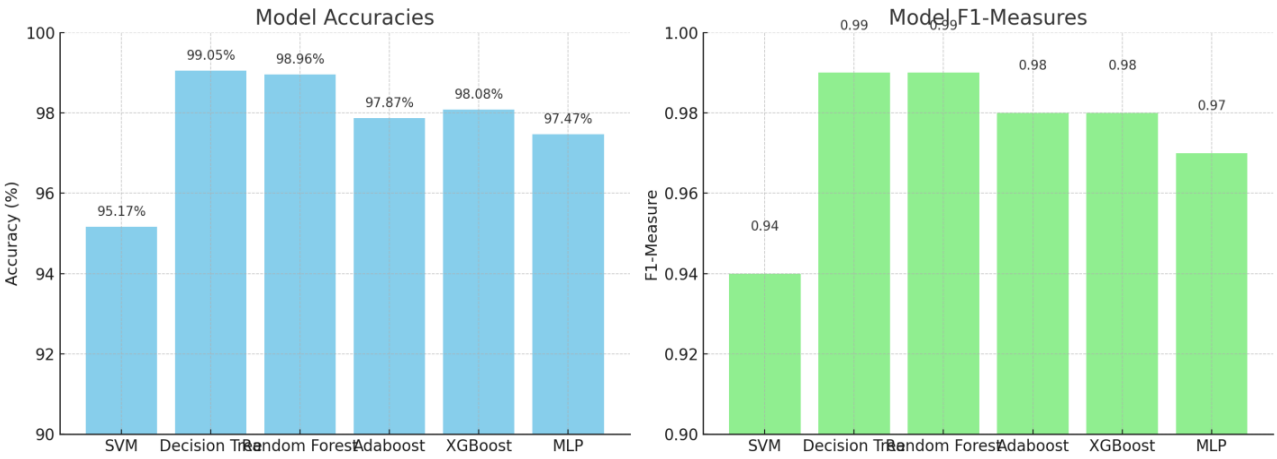

Рис. 2. Порівняльний аналіз продуктивнсоті та збалансованої оцінки ефективності моделі

Для оцінки алгоритму KNN (табл. 2, рис. 3) було протестовано декілька моделей з різними значеннями K (від 2 до 9). Найвища Accuracy на рівні 95,58, була досягнута при значенні K, рівному 7. Вибрані метрики для навчання моделі виявилися дуже релевантними, ефективно відрізняючи шаблони звичайного трафіку від зловмисного.

Це свідчить про те, що основні ознаки були достатньо виразними, а межі рішень могли бути ефективно відображені за допомогою ієрархічної структури DT.

Таблиця 2

Результати роботи моделей KNN

Модель	Recall	Precision	Accuracy	F1 measure
2-NN	95	93	94,51	94
3-NN	95	95	95,47	95
4-NN	95	94	95,12	94
5-NN	95	95	95,56	95
6-NN	95	94	95,31	95
7-NN	**94**	**95**	**95,58**	**95**
8-NN	95	95	95,48	95
9-NN	94	95	95,57	95

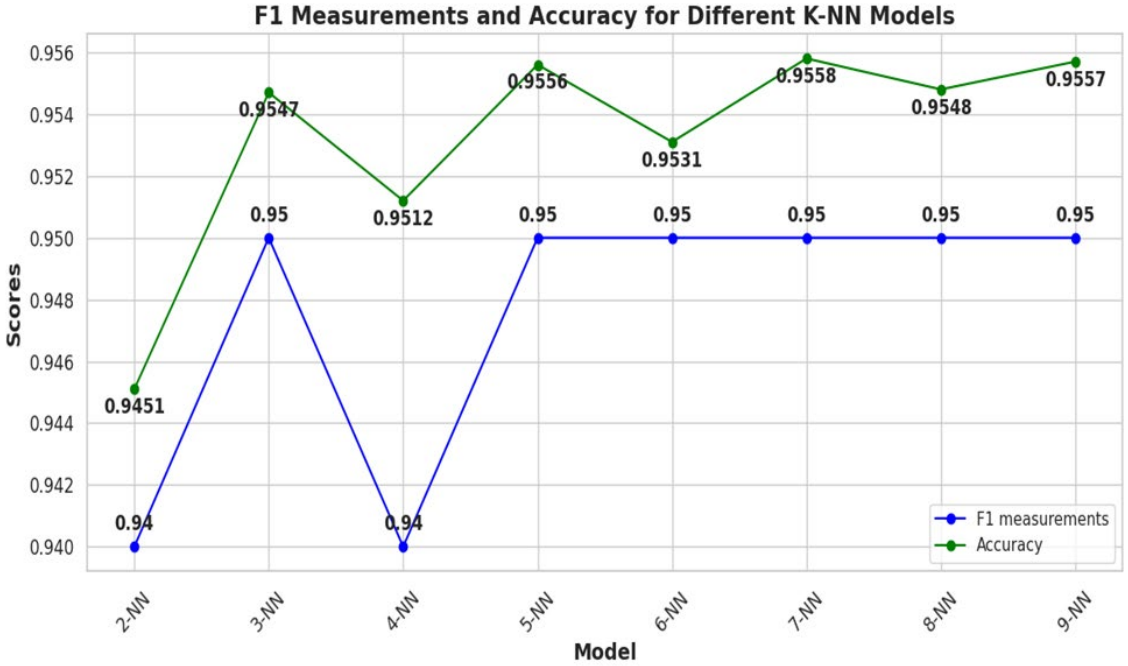

Рис. 3. Відновлення KNN після виявлення атаки

Хоча модель SVM досягла Recall на рівні 93 %, що свідчить про її здатність розпізнавати велику частку атак (істинних спрацьовувань), її Accuracy була відносно нижчою порівняно з іншими моделями.

Порівняно з попередніми роботами дослідників [2], результати яких показали Accuracy на рівні 85,56 % для DT та 81,34 % для штучних нейронних мереж (ANN), запропонована система досягла значно вищої Accuracy на рівні 99,05 % для DT та 97,47 % для MLP.

Це покращення підкреслює релевантність та розпізнавальну здатність ознак у наборі даних CICIDS2017, що свідчить про те, що відбір ознак може бути необов’язковим і може потенційно знизити Precision.

Продуктивність моделей DL

Модель Deep MLP продемонструвала відмінні результати виявлення мережевих атак, як показано в таблиці 3 та на рис. 4. Завдяки оптимізатору Adam і співвідношенню об’єму навчальної та тестової вибірки 80:20 вона досягла високої Accuracy на рівні 98,44 % та високої

збалансованої оцінки ефективності моделі на рівні 98 %, що демонструє свою здатність ефективно виявляти атаки, мінімізувавши при цьому помилкові класифікації.

Оптимізатор Adam мав більшу продуктивність ніж оптимізатор Stochastic Gradient Descent, що сприяло високим результатам Deep MLP. Співвідношення 80:20 між тренуванням і тестом відповідає розподілу звичайного і зловмисного трафіку в реальних системах, що дозволяє моделі добре узагальнювати приховані дані.

Таблиця 3

Результати роботи оптимізаторів SGD та ADAM

Optimizer	Train: Test	Recall	Precision	Accuracy	F1-measure	AUC
SGD	90:10	98	98	98,19	98	97,6
SGD	80:20	97	98	97,68	97	97,1
SGD	70:30	98	98	97,99	98	97,5
SGD	60:40	97	98	97,82	97	97,2
ADAM	90:10	98	98	98,11	98	97,5
ADAM	80:20	98	98	98,44	98	98,1
ADAM	70:30	98	98	98,36	98	98
ADAM	60:40	98	98	98,35	98	97,9

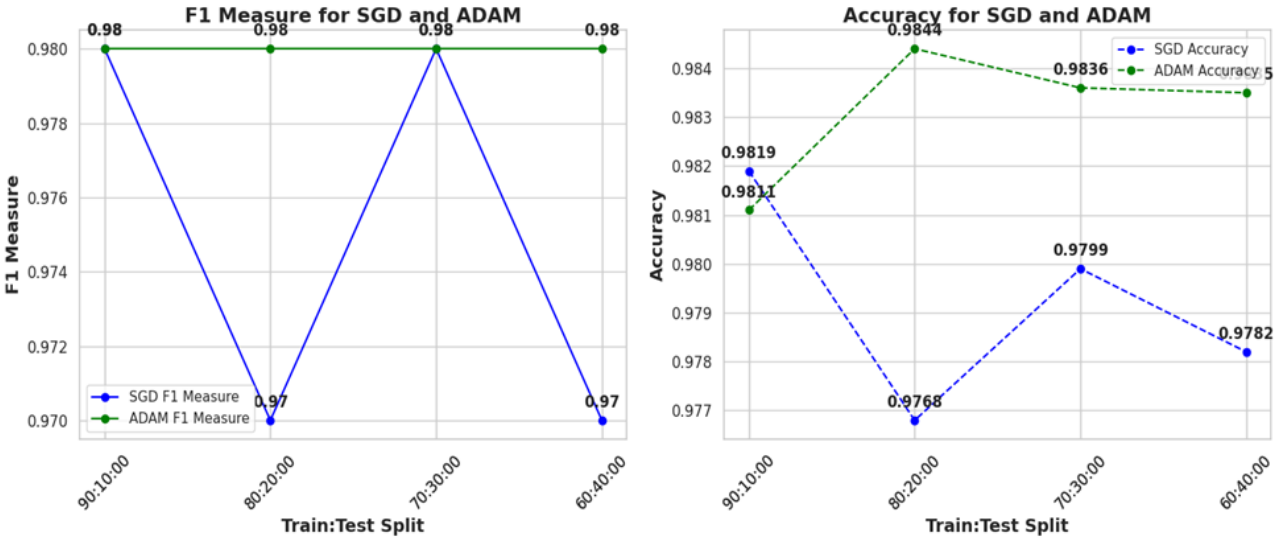

Рис. 4. Порівняння точності та F1-оцінки Deep MLP

Ефективність моделей виявлення атак

У таблицях 4–8 представлено ефективність чисельних моделей ML, EL та DL у виявленні дев'яти різних типів мережевих атак.

Далі в таблиці 4 представлена ефективність різних моделей ML, EL та DL у виявленні атак Analysis. Модель RF демонструє належні показники в усіх метриках, а збалансована оцінка ефективності моделі в межах 24 %. З іншого боку, модель XGBoost має найвищу Accuracy на рівні 72 %, але її Recall та Precision значно нижчі, що свідчить про те, що вона не може послідовно ідентифікувати всі справжні тривоги. Модель AdaBoost демонструє погані результати, з усіма метриками на рівні або нижче 5 %, що свідчить про те, що вона значно гірше справляється з цим типом атак. Моделі MLP та DT мають схожу картину, з дещо вищою Precision, але нижчими показниками Recall та Accuracy. Нарешті, модель 7-NN досягає відносного балансу, з показниками Recall та збалансованої оцінки ефективності моделі до

17 % та 25 % відповідно, забезпечуючи більш середню Accuracy порівняно з іншими оціненими моделями.

Таблиця 4

Analysis атаки та ідентифікація результатів

Model	**Recall**	**Precision**	**Accuracy**	**F1-measure**
DT	13	51	13	20
7-NN	17	47	17	25
RF	24	25	24	24
AdaBoost	2	5	2	2
XGBoost	20	72	20	31
MLP	13	64	13	22

У таблиці 5 представлена ефективність різних моделей ML у виявленні DoS-атак. Модель RF має збалансовані метрики, з Precision на рівні 38 % та Accuracy на рівні 36 %. Дана модель має достатню здатність точно ідентифікувати позитивні випадки без великої кількості хибнонегативних результатів. Модель XGBoost має найвищу Precision на рівні 42 %, але немає належного рівня до Recall та Accuracy. Це вказує на те, що модель не здатна ідентифікувати більшість справжніх позитивних DoS-атак. Моделі AdaBoost та MLP мають обмежені значення у виявленні DoS-атак, що призводить до низьких показників за всіма метриками. Це свідчить про більший відсоток хибних спрацьовувань і загальну неефективність у цьому випадку. Модель DT майже неефективна, з усіма метриками на рівні 1 %, що свідчить про її недостатню ефективність. Модель 7-NN постійно досягає показників на рівні 30 %, що свідчить про помірний рівень Accuracy, Precision та збалансованості її можливостей виявлення.

Таблиця 5

Результати ідентифікації DoS-атак

Model	Recall	Precision	Accuracy	F1-measure
DT	1	1	1	1
7-NN	30	30	30	30
RF	36	38	36	37
AdaBoost	3	10	3	5
XGBoost	4	42	4	6
MLP	10	40	10	15

У таблиці 6 представлена ефективність різних моделей у виявленні Exploits, які використовують слабкі місця в мережах. Модель RF серед всіх запропонованих моделей, має високу ефективність з Precision на рівні 75 % та Accuracy на рівні 80 %. Це свідчить про те, що вона надійно виявляє атаки з використанням Exploits, не генеруючи при цьому багато хибних спрацьовувань. XGBoost має нижчу Precision на рівні 62 %, але перевершує її за метриками Recall та Accuracy на рівні 94 %, що робить її високоефективною у виявленні більшості істинно позитивних випадків, що має вирішальне значення в певних сценаріях. Модель AdaBoost, незважаючи на Precision на рівні 57 % та Accuracy на рівні 36 %, демонструє помірну працездатність, але має деякі проблеми із загальним представленням, про що свідчать її нижчі показники Recall та збалансованої оцінки ефективності. Модель MLP має високу метрику Recall та Accuracy на рівні 88 %. Це свідчить про те, що вона ефективно виявляє атаки з використанням Exploits. Нарешті, модель DT, вона хоч і має нижчу Precision на рівні 54 %, але демонструє дуже високі показники у Recall та Accuracy на рівні 91 %. Це дає змогу стверджувати, що вона може ідентифікувати майже всі атаки з використанням Exploits, але за рахунок більшої кількості помилкових спрацьовувань.

Таблиця 6

Результати ідентифікації атак з використанням Exploits

Model	Recall	Precision	Accuracy	F1-measure
DT	91	54	91	68
7-NN	74	63	74	68
RF	80	75	80	77
AdaBoost	36	57	36	44
XGBoost	94	62	94	74
MLP	88	62	88	73

У таблиці 7 представлено ефективності різних моделей у виявленні Generic атак, які включають широкий спектр методів атаки без конкретного націлювання на вразливості програмного забезпечення. Моделі RF, XGBoost та MLP демонструють високу Precision та Accuracy, а їхні результати наближаються до ідеальних. Модель RF має Precision на рівні 98 % та Accuracy на рівні 97 %, що свідчить про її здатність точно ідентифікувати Generic атаки, мінімізуючи при цьому помилкові спрацьовування. Модель XGBoost та MLP дещо перевершує модель RF за Precision на рівні 99 %, але має однаковий Recall та Accuracy, демонструючи її ефективність класифікувати реальні випадки Generic атак належним чином та без суттєвих пропусків. Модель AdaBoost майже неефективна, з усіма метриками на рівні 1–2 %, що свідчить про її недостатню ефективність.

Таблиця 7

Результати ідентифікації Generic атак

Model	Recall	Precision	Accuracy	F1-measure
RF	97	98	97	98
AdaBoost	1	4	1	2
XGBoost	97	99	97	98
MLP	97	99	97	98

У таблиці 8 представлена ефективність різних моделей у виявленні атак Worm, Backdoor та Shellcode, отриману за допомогою Accuracy та збалансованої оцінки ефективності. Серед представлених моделей, SVM демонструє найвищу Accuracy для Worm і Shellcode на рівні 38 % та 39 %, що свідчить про його здатність розрізняти різні типи атак. Однак збалансована оцінка ефективності для SVM, для Worm і Shellcode, є відносно низька, що вказує на потенційні труднощі в досягненні балансу між Precision та Recall. AdaBoost демонструє Accuracy із Precision на рівні від 27 % до 65 % і з відносно вищою збалансованою оцінкою ефективності у всіх категоріях порівняно з іншими моделями. І навпаки, такі моделі, як MLP, демонструють нижчі показники Precision на рівні від 13 % до 50 %, і відносно послідовну збалансовану оцінку ефективності для різних типів атак. Загалом, хоча деякі моделі демонструють кращу Precision, їхні збалансовані оцінки ефективності потребують подальшого вдосконалення, щоб мати змогу ефективно розпізнавати кожну категорію атаки. Ці результати підкреслюють важливість постійної оптимізації для підвищення ефективності систем виявлення кіберзагроз.

Таблиця 8

Результати ідентифікації атак Worm, Backdoor and Shellcode

Model	Accuracy			F1-measure		
	Worm	Backdoor	Shellcode	Worm	Backdoor	Shellcode
DT	26	29	37	41	27	48
SVM	38	30	39	25	48	37
RF	35	54	53	19	43	32
AdaBoost	27	53	65	48	60	59
XGBoost	46	45	24	31	34	43
MLP	13	42	50	27	30	48

Висновки з даного дослідження і перспективи подальших досліджень у даному напрямку. У представленому дослідженні з використанням набору даних CICIDS2017, автори ознайомились з ефективністю численних моделей ML і DL у виявленні мережевих атак (вторгнень). Результати показали чудову продуктивність на рівні 99,05 класичної моделі ML у вигляді класифікатора DT порівняно з такими підходами, як XGBoost, RF та AdaBoost. Модель Deep MLP, навчена за допомогою оптимізатора Adam, досягла продуктивності на рівні 98,44 при співвідношенні тренувань і тестів як 80:20.

Моделі показали успішну ідентифікацію певних типів атак, таких як Exploits та Generic атаки. Однак, вони зіткнулися з труднощами в точному виявленні таких категорій, як DoS-атаки, Worm, Backdoor та Shellcode. Це вказує на необхідність вдосконалення здатності моделей класифікувати ці конкретні типи атак. Підвищення продуктивносиі моделей у цих категоріях має вирішальне значення для посилення заходів кібербезпеки та забезпечення комплексних можливостей виявлення загроз. Результати досліджень мають важливе значення для захисту критично важливих мереж від кібератак. Виявлення вторгнень у мережах можна покращити, використовуючи моделі ML та DL, а також нових методологій, таких як навчання згорткових нейронних мереж (CNN) безпосередньо на захоплених пакетах мережевого трафіку або візуальних зображеннях. Ці рішення можуть ефективно виявляти як відомі, так і невідомі методи атак, забезпечуючи повний захист від складних кіберзагроз, з якими стикаються мережі. У майбутньому більш складні алгоритми ML можуть бути використані для виявлення кібератак з більшою точністю та ширшим класом кібератак.

Подальші дослідження виявлення вторгнень у мережі повинні бути зосереджені на покращенні стійкості моделі проти складних типів атак, таких як DoS-атаки і Backdoors, за допомогою тестування та розширення бази даних. Інтеграція сукупності цих методів з передовими архітектурами DL, таких як CNN та RNN (рекурентна нейронна мережа), може підвищити точність виявлення. Виявлення вторгнень у режимі реального часу можна покращити, використовуючи фреймворки потокової обробки та периферійні обчислення. Крім того, зрозумілі методи штучного інтелекту забезпечать прозорість і довіру до мереж, що має вирішальне значення для середовищ із високими ризиками.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Li Y., Liu Q. A comprehensive review study of cyber-attacks and cyber security; Emerging trends and recent developments. *Energy Reports*. 2021. URL: https://doi.org/10.1016/j.egyr.2021.08.126 (date of access: 03.05.2024).

2. George A. S., George A. H. & Baskar T. Digitally Immune Systems: Building Robust Defences in the Age of Cyber Threats. *Partners Universal International Innovation Journal*. 2023. Vol. 1. № 4. P. 155–172. DOI: 10.5281/zenodo.8274514.

3. Explainable Intrusion Detection for Cyber Defences in the Internet of Things: Opportunities and Solutions / N. Moustafa et al. *IEEE Communications Surveys & Tutorials*. 2023. P. 1. URL: https://doi.org/10.1109/comst.2023.3280465 (date of access: 03.05.2024).

4. A Review of Intrusion Detection Systems Using Machine and Deep Learning in Internet of Things: Challenges, Solutions and Future Directions / J. Asharf et al. *Electronics*. 2020. Vol. 9, no. 7. P. 1177. URL: https://doi.org/10.3390/electronics9071177 (date of access: 03.05.2024).

5. An Experimental Analysis of Attack Classification Using Machine Learning in IoT Networks / A. Churcher et al. *Sensors*. 2021. Vol. 21, no. 2. P. 446. URL: https://doi.org/10.3390/s21020446 (date of access: 03.05.2024).

6. Evaluating Deep Learning Based Network Intrusion Detection System in Adversarial Environment / Y. Peng et al. *2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC)*, Beijing, China, 12–14 July 2019. 2019. URL: https://doi.org/10.1109/iceiec.2019.8784514 (date of access: 03.05.2024).

7. Intrusion Detection and Prevention Systems: An Updated Review / N. A. Azeez et al. *Data Management, Analytics and Innovation*. Singapore, 2019. P. 685–696. URL: https://doi.org/10.1007/978-981-32-9949-8_48 (date of access: 03.05.2024).

8. A Survey of Intrusion Detection Systems Leveraging Host Data / R. A. Bridges et al. *ACM Computing Surveys*. 2020. Vol. 52, no. 6. P. 1–35. URL: https://doi.org/10.1145/3344382 (date of access: 03.05.2024).

9. Parizad A., Hatziadoniu C. Cyber-Attack Detection Using Principal Component Analysis and Noisy Clustering Algorithms: A Collaborative Machine Learning-Based Framework. *IEEE Transactions on Smart Grid*. 2022. Vol. 13, no. 6. P. 4848-4861. URL: https://doi.org/10.1109/tsg.2022.3176311 (date of access: 03.05.2024).

10. Eliyan L. F., Di Pietro R. DoS and DDoS attacks in Software Defined Networks: A survey of existing solutions and research challenges. *Future Generation Computer Systems*. 2021. Vol. 122. P. 149–171. URL: https://doi.org/10.1016/j.future.2021.03.011 (date of access: 03.05.2024).

11. Alghawazi M., Alghazzawi D., Alarifi S. Detection of SQL Injection Attack Using Machine Learning Techniques: A Systematic Literature Review. *Journal of Cybersecurity and Privacy*. 2022. Vol. 2, no. 4. P. 764–777. URL: https://doi.org/10.3390/jcp2040039 (date of access: 03.05.2024).

12. SSH and FTP brute-force Attacks Detection in Computer Networks: LSTM and Machine Learning Approaches / M. D. Hossain et al. *2020 5th International Conference on Computer and Communication Systems (ICCCS)*, Shanghai, China, 15–18 May 2020. 2020. URL: https://doi.org/10.1109/icccs49078.2020.9118459 (date of access: 03.05.2024).

13. Thankappan M., Rifà-Pous H., Garrigues C. Multi-Channel Man-in-the-Middle Attacks Against Protected Wi-Fi Networks: A State of the Art Review. *Expert Systems with Applications*. 2022. Vol. 210. P. 118401. URL: https://doi.org/10.1016/j.eswa.2022.118401 (date of access: 03.05.2024).

14. Almulla K. Cyber-attack detection in network traffic using machine learning. 2022. URL: https://repository.rit.edu/cgi/viewcontent.cgi?article=12453&context=theses (date of access: 03.05.2024).

15. Deep Learning Approach for Intelligent Intrusion Detection System / R. Vinayakumar et al. *IEEE Access*. 2019. Vol. 7. P. 41525–41550. URL: https://doi.org/10.1109/access.2019.2895334 (date of access: 03.05.2024).

16. BAT: Deep Learning Methods on Network Intrusion Detection Using NSL-KDD Dataset / T. Su et al. *IEEE Access*. 2020. Vol. 8. P. 29575–29585. URL: https://doi.org/10.1109/access.2020.2972627 (date of access: 03.05.2024).

17. Ding Y., Zhai Y. Intrusion Detection System for NSL-KDD Dataset Using Convolutional Neural Networks. *CSAI '18: Proceedings of the 2018 2nd International Conference on Computer Science and Artificial Intelligence,* Shenzhen, China, 8–10 December 2018. New York, USA, ACM Press, 2018. P. 81–85. URL: https://doi.org/10.1145/3297156.3297230 (date of access: 03.05.2024).

18. Meena G., Choudhary R. R. A review paper on IDS classification using KDD 99 and NSL KDD dataset in WEKA. *2017 International Conference on Computer, Communications and Electronics (Comptelix)*, Jaipur, India, 1–2 July 2017. 2017. URL: https://doi.org/10.1109/comptelix.2017.8004032 (date of access: 03.05.2024).

19. Benchmarking datasets for Anomaly-based Network Intrusion Detection: KDD CUP 99 alternatives / A. Divekar et al. *2018 IEEE 3rd International Conference on Computing, Communication and Security (ICCCS)*, Kathmandu, 25–27 October 2018. 2018. URL: https://doi.org/10.1109/cccs.2018.8586840 (date of access: 03.05.2024).

20. A Deep Learning Approach for Network Intrusion Detection Based on NSL-KDD Dataset / C. Zhang et al. *2019 IEEE 13th International Conference on Anti-counterfeiting, Security, and Identification (ASID)*, Xiamen, China, 25–27 October 2019. IEEE. P. 41–45. URL: https://doi.org/10.1109/icasid.2019.8925239 (date of access: 04.05.2024).

21. CICIDS2017. URL: https://www.researchgate.net/figure/Description-of-files-containing-CICIDS2017-dataset_tbl1_329045441 (date of access: 04.05.2024).