КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 3 (27), 2025
ISSN 2663 - 4023
DOI 10.28925/2663-4023.2025.27.711
УДК 004.8

Журавчак Анастасія Юріївна
аспірант, асистент кафедри захисту інформації
Національний Університет «Львівська Політехніка», Львів, Україна
ORCID ID: 0000-0002-8196-7963
anastasiia.y.tolkachova@lpnu.ua

Піскозуб Андріян Збігнєвич
кандидат технічних наук, доцент кафедри захисту інформації
Національний Університет «Львівська Політехніка», Львів, Україна
ORCID ID: 0000-0002-3582-2835
andriian.z.piskozub@lpnu.ua

АНАЛІЗ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ АВТОМАТИЗАЦІЇ ТЕСТУВАННЯ НА ПРОНИКНЕННЯ

Анотація. Автоматизація тестування на проникнення за допомогою методів машинного навчання є однією з найбільш перспективних галузей у сучасній кібербезпеці. Традиційний підхід до тестування на проникнення вимагає значних ресурсів, зокрема фінансових, а також залучення висококваліфікованих фахівців, здатних проводити комплексну оцінку безпеки систем. Такий підхід не завжди може забезпечити достатню швидкість виявлення нових загроз, особливо в умовах постійно зростаючої складності кібератак і великої кількості вразливостей. Впровадження методів машинного навчання у процес пентесту дозволяє створювати гнучкі, адаптивні системи, які можуть не лише автоматизувати рутинні завдання, але й підвищити точність та ефективність виявлення вразливостей. У статті здійснено огляд ключових алгоритмів машинного навчання, які можуть бути використані для автоматизації тестування на проникнення, включаючи метод опорних векторів, випадковий ліс, наївний Байєс, дерева рішень, а також методи навчання з підкріпленням. Кожен з цих алгоритмів пропонує певні переваги у контексті аналізу вразливостей, класифікації загроз та визначення пріоритетів для виправлення критичних проблем безпеки. Окрему увагу приділено ролі великих мовних моделей у процесі автоматизації. Вони можуть виконувати завдання аналізу логів, класифікації загроз, генерування звітів і навіть надавати рекомендації для усунення виявлених вразливостей. Такі моделі дозволяють значно підвищити продуктивність фахівців, виконуючи рутинні завдання автоматично, що особливо корисно в умовах інтеграції з CI/CD процесами. Водночас застосування LLM має певні обмеження, зокрема залежність від актуальних даних та високу вартість обчислень. У статті також розглядаються виклики та обмеження впровадження алгоритмів машинного навчання у процес пентесту, такі як потреба у великій кількості якісних даних для тренування моделей, високі обчислювальні ресурси, а також ризики, пов’язані з можливими помилковими спрацьовуваннями. Результати дослідження демонструють, що алгоритми машинного навчання мають значний потенціал для підвищення ефективності автоматизованого тестування на проникнення, особливо у великих інфраструктурах з численними вразливими точками.

Ключові слова: машинне навчання; тестування на проникнення; глибоке навчання; великі мовні моделі; дерева рішень; SVM.

ВСТУП

Традиційний процес тестування на проникнення відомий своїми значними фінансовими зобов’язаннями та потребою у високому рівні експертизи. Він вимагає мати в штаті або за контрактом спеціалістів, які впродовж певного часу проведуть оцінку стану безпеки. Однак, зростання кількості кіберзагроз та збільшення складності атак

створюють нові виклики для забезпечення своєчасного й ефективного захисту. У зв'язку з цим автоматизація пентесту стала пріоритетом для багатьох компаній та організацій.

Методи машинного навчання відкривають нові можливості для автоматизації тестування на проникнення. Вони дозволяють не тільки ідентифікувати відомі вразливості, але й виявляти нові загрози шляхом аналізу поведінки та ознак аномальної активності. Інтеграція машинного навчання у процес пентесту надає змогу створювати більш гнучкі та адаптивні системи. Вони здатні ефективно моделювати сценарії атак та розпізнавати складні шаблони загроз. Такі системи можуть стати цінним інструментом для фахівців з кібербезпеки, надаючи їм можливість фокусуватись на критичних аспектах аналізу та оцінки ризиків.

Постановка проблеми. Традиційні методи тестування є трудомісткими та не завжди встигають за темпом нових загроз. Використання машинного навчання для автоматизації тестування на проникнення дозволяє швидше виявляти вразливості. Викликом є вибір найбільш ефективного методу машинного навчання для тестування на проникнення.

Аналіз останніх досліджень і публікацій. У сфері автоматизації тестування на проникнення за допомогою методів машинного навчання спостерігається значний науковий інтерес. Дослідники зосереджуються на розробці інтелектуальних систем, здатних ефективно виявляти та експлуатувати вразливості в інформаційних системах.

В роботі [1] представлено підхід до виявлення вразливостей у програмному забезпеченні за допомогою великих мовних моделей (LLM). Авторами запропоновано метод статичного аналізу коду, який інтегрує можливості LLM для автоматичного виявлення потенційних загроз. Система демонструє ефективність у виявленні складних уразливостей, які важко виявити за допомогою традиційних інструментів аналізу. Показано, що використання LLM дозволяє зменшити кількість хибнопозитивних результатів. Однак відзначено, що масштабованість рішення та залежність від якості вихідних даних залишаються викликами для широкого впровадження цього підходу.

В роботі [2] представлені результати досліджень щодо використання машинного навчання для автоматизованого проведення тестувань на проникнення. Авторами запропоновано метод, який поєднує машинне навчання з традиційними підходами до оцінки безпеки. Система демонструє здатність автоматично ідентифікувати слабкі місця у захисті, що суттєво зменшує час, необхідний для аналізу. Зроблено акцент на використанні моделей машинного навчання для оцінки безпеки сучасних веб-додатків. Однак залишаються відкритими питання щодо точності виявлення складних вразливостей та адаптації методів до різних середовищ безпеки.

В роботі [3] представлено підхід до виявлення вразливостей у програмному забезпеченні за допомогою великих мовних моделей. Автори досліджують можливість використання LLM для аналізу коду з метою автоматичного виявлення слабких місць у програмному забезпеченні. Запропоновано систему, яка об'єднує методи обробки природної мови з інструментами статичного аналізу для виявлення як типових, так і складних вразливостей. Важливим результатом роботи є зменшення кількості хибнопозитивних спрацювань, що підвищує точність аналізу. Проте автор відзначає залежність ефективності системи від навчальних даних, а також обмеження в масштабованості, що потребує подальших досліджень.

Підсумовуючи ці дослідження, можна зазначити, що існуючі методи виявлення вразливостей у програмному забезпеченні на основі традиційних підходів мають свої обмеження. Зокрема, основною проблемою є залежність від ручних перевірок, специфікацій безпеки, а також обмежена здатність до масштабного аналізу великих

кодових баз. Таким чином, можна помітити існуючу необхідність розробки та впровадження інтегрованих рішень, які поєднують можливості великих мовних моделей із традиційними методами статичного та динамічного аналізу. Це дозволить підвищити точність, зменшити кількість хибнопозитивних спрацювань і автоматизувати виявлення складних вразливостей у різних програмних середовищах.

Метою статті є висвітлення різних методів машинного навчання для тестування на проникнення.

МЕТОДИКА ДОСЛІДЖЕННЯ

Для порівняння взяті вже існуючі методи машинного навчання для проведення тестування на проникнення. Проаналізовано їхні можливості та сильні сторони.

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

У сучасних умовах кібербезпеки методи машинного навчання стають важливим інструментом для автоматизації тестування на проникнення. Вони дозволяють більш ефективно й точно виявляти загрози та оцінювати вразливості. Вибір відповідного алгоритму машинного навчання є критичним для побудови надійної системи автоматизованого пентесту, здатної визначати та класифікувати вразливості. У цьому розділі розглянуто основні алгоритми, які можуть підвищити ефективність тестування на проникнення. Зокрема метод опорних векторів (SVM), випадковий ліс (Random Forest), контрольований інтелектуальний аналіз даних, наївний Байєс, дерева рішень, метод найменших квадратів для SVM (LS-SVM), навчання з підкріпленням (Reinforcement Learning) та великі мовні моделі (LLM). Кожен із цих методів пропонує специфічні переваги для вирішення задач аналізу та виявлення вразливостей, забезпечуючи тестувальників гнучкими та ефективними засобами для підвищення точності й оперативності процесу пентесту.

Метод опорних векторів може застосовуватися для автоматизації різних аспектів оцінки вразливостей і виявлення загроз. SVM може класифікувати виявлені вразливості за ступенем серйозності або можливості використання [4]. Це може допомогти тестувальникам визначити пріоритети вразливостей.

За допомогою Random Forest вразливості можна класифікувати на основі їх серйозності, можливості використання або впливу. Тренуючи модель на раніше відомих вразливостях, вона може допомогти пентестерам визначити пріоритетність проблем з високим ризиком. Критичні вразливості будуть усунені в першу чергу. Random Forest дає уявлення про те, які ознаки (наприклад, мережеві параметри, заголовки запитів, властивості файлів) є найбільш передбачуваними для інцидентів безпеки [5]. Це допомагає пентестерам зосередитися на найбільш релевантних даних. Інтегруючи моделі випадкових лісів в інструменти пентесту тестувальники можуть більш ефективно розставляти пріоритети вразливостей і зменшити кількість помилкових спрацьовувань, тим самим роблячи процес більш надійним і глибоким.

Контрольований інтелектуальний аналіз (Supervised data mining) даних може відігравати важливу роль у тестуванні на проникнення, використовуючи марковані дані для навчання моделей, які передбачають і виявляють загрози безпеці [6]. Навчаючись на історичних даних про вразливості, контрольовані моделі можуть передбачати ймовірність появи нових вразливостей в системах або додатках. Моделі класифікації

також можуть класифікувати вразливості на основі рівня ризику, що допомагає визначити пріоритетність найбільш важливих проблем для виправлення під час пентесту. Контрольовані моделі можуть навчатися на відомих шаблонах атак (наприклад, SQL-ін'єкція, XSS) і виявляти подібні шаблони в сценаріях пентесту. Цей підхід особливо корисний для виявлення вразливостей веб-додатків шляхом зіставлення відомих сигнатур атак з результатами тестів. Аналізуючи марковані дані про типову поведінку користувачів, керовані моделі можуть виявити аномалії в активності облікових записів. Такі як незвичний час входу в систему або місцезнаходження. Це може допомогти пентестерам оцінити ефективність програми у виявленні захоплення акаунтів або інсайдерських загроз.

Наївний Байєс, алгоритм машинного навчання, може бути корисним у тестуванні на проникнення, особливо для класифікації та виявлення загроз безпеці на основі історичних даних [7]. Наївний Байєс може бути навчений на наборах даних поширених сигнатур веб-атак, таких як SQL-ін'єкції або шаблони міжсайтового скриптингу (XSS), щоб відмічати подібні спроби атак в режимі реального часу. Це корисно для виявлення типових веб-вразливостей у веб-додатках. Аналізуючи особливості активності користувача (наприклад, час входу в систему, частоту доступу), Naive Bayes може виявити відхилення від типової поведінки. Це допомагає перевірити здатність програми виявляти внутрішні загрози або підозрілі дії користувачів. Хоча наївний Байєс передбачає незалежність від ознак, що не завжди відповідає дійсності в контексті безпеки, він залишається ефективним і інтерпретованим, що робить його корисним алгоритмом для легких завдань класифікації в реальному часі при тестуванні. Він особливо добре підходить для сценаріїв, де важливе швидке ймовірнісне виявлення.

Дерева рішень ефективні для відображення потенційних шляхів атак в мережі [8]. Моделюючи залежності безпеки та можливі вектори атак, вони допомагають пентестерам визначити найбільш ймовірні та ефективні шляхи атак, імітуючи, як зловмисник може рухатися через систему. Дерева рішень можуть класифікувати дії користувачів на основі шаблонів поведінки, таких як час входу в систему та тривалість сеансу. Вони можуть виявляти аномалії, які можуть вказувати на компрометацію облікового запису, допомагаючи пентестерам оцінити здатність програми розпізнавати внутрішні загрози. Вивчаючи відомі шаблони веб-атак (наприклад, SQL-ін'єкції, XSS), дерева рішень можуть допомогти у виявленні та класифікації вхідних веб-запитів у реальному часі.

Метод найменших квадратів опорних векторів (LS-SVM) — це адаптація традиційного методу опорних векторів (SVM), який можна застосовувати в тестуванні на проникнення для покращеного виявлення загроз та аналізу вразливостей [9]. LS-SVM може класифікувати вразливості на основі серйозності або типу, допомагаючи пентестерам визначити пріоритетність вразливостей з високим ризиком. Цей підхід цінний тим, що дозволяє зосередити зусилля з тестування на областях, де вплив або ймовірність експлуатації є найвищою, тим самим підвищуючи ефективність тестування. LS-SVM можна використовувати для виявлення веб-атак, таких як SQL-ін'єкції, міжсайтовий скриптинг та інші поширені вектори атак. LS-SVM пропонує перевагу обчислювальної ефективності завдяки підходу найменших квадратів, який спрощує оптимізацію, роблячи її в деяких випадках швидшою порівняно з традиційним SVM. Ця ефективність особливо корисна в задачах тестування в реальному часі або масштабних завданнях, де швидка класифікація загроз і вразливостей має важливе значення.

Навчання з підкріпленням може бути потужним інструментом у тестуванні на проникнення, дозволяючи системам навчатися та адаптуватися до складних середовищ

безпеки. RL може імітувати поведінку зловмисника, вивчаючи оптимальну послідовність дій для використання вразливостей і досягнення конкретних цілей в цільовій системі. Методом проб і помилок агент RL може досліджувати шляхи атаки, вивчаючи, як переміщатися через мережу або додаток, щоб досягти цінних активів. Моделі RL можуть оптимізувати стратегії сканування вразливостей шляхом коригування параметрів на основі зворотного зв'язку системи, таких як час відгуку або уникнення виявлення [10]. Цей динамічний підхід дозволяє агенту RL ефективніше виявляти вразливості, мінімізуючи ймовірність виявлення, імітуючи складні тактики зловмисників. Вивчаючи різні підходи і спостерігаючи за реакцією мережі, агенти RL можуть виявити слабкі місця в інфраструктурі безпеки, які можуть дозволити реальним зловмисникам обійти захист. Навчання з підкріпленням може допомогти визначити пріоритети ресурсів, визначаючи найбільш ефективні області для атак у великій мережі. RL можна використовувати для динамічного тестування веб-додатків, адаптуючи тести на основі зворотного зв'язку від відповідей програми. Це включає визначення можливих точок входу, створення цільових корисних навантажень і дослідження функціональності програми в режимі реального часу для виявлення прихованих вразливостей. RL можна застосовувати для моделювання та оптимізації фішингових атак, дізнаючись, які типи повідомлень або тактики соціальної інженерії мають більше шансів на успіх. Це допомагає організаціям зрозуміти свою вразливість до таких атак і вдосконалити програми навчання та підвищення обізнаності. Здатність Reinforcement Learning до ітеративного навчання та адаптації на основі зворотного зв'язку робить його добре придатним для складних, мінливих сценаріїв тестування.

Використання LLM для автоматизації пентесту дозволяє пришвидшити і спростити виявлення вразливостей. Модель може аналізувати логи, класифікувати загрози, генерувати коди експлойтів, створювати звіти й давати рекомендації. Це особливо корисно для рутинних задач та інтеграції з CI/CD. Проте виклики включають потребу в актуальних даних та високу вартість обчислень, тому LLM поки не може замінити ручне тестування, але здатна значно підвищити його ефективність [11].

Використання методів машинного навчання для автоматизації тестування на проникнення має значний потенціал для підвищення ефективності та точності процесу виявлення вразливостей. Різні алгоритми, такі як метод опорних векторів, Random Forest, наївний Байєс, дерева рішень, навчання з підкріпленням, LS-SVM та великі мовні моделі, пропонують унікальні підходи до класифікації загроз, визначення пріоритетів і аналізу даних, що дозволяє зробити процес пентесту більш гнучким і адаптивним.

Кожен метод має свої переваги та обмеження, і їх вибір залежить від конкретних вимог і характеристик системи, яку тестують. Інтеграція алгоритмів машинного навчання в інструменти та робочі процеси пентестингу дозволяє тестувальникам зосередитися на критичних загрозах, знижуючи ризики та підвищуючи швидкість реакції на нові виклики кібербезпеки. Однак для ефективної роботи таких моделей необхідно мати достатню кількість якісних даних і значні обчислювальні ресурси.

У цілому, машинне навчання має потенціал трансформувати процес тестування на проникнення, зробивши його більш надійним і масштабованим, проте подальші дослідження і вдосконалення алгоритмів є необхідними для подолання поточних обмежень і забезпечення їхньої максимальної ефективності в реальних умовах.

На рис. 1 наведено [12] кількісну оцінку впливу методів машинного навчання та штучного інтелекту на ефективність тестування на проникнення у порівнянні з традиційними підходами. Використання ML і AI значно зменшує середній час виявлення вразливостей з 48 годин (традиційні методи) до 6 годин, що вказує на суттєве

прискорення аналізу. ML і AI знижують показник хибнопозитивних результатів з 15% до 5% і хибнонегативних з 10% до 3%, що підвищує точність тестування. Вони зменшують середній час генерації атак з 72 годин до 12 годин, дозволяючи швидше реагувати на потенційні загрози. Успішність експлуатації вразливостей для AI-методів становить 90%, що перевищує традиційні показники 60%. AI зменшує середній час виявлення порушень з 168 годин до 72 годин, що дозволяє швидше реагувати на інциденти.

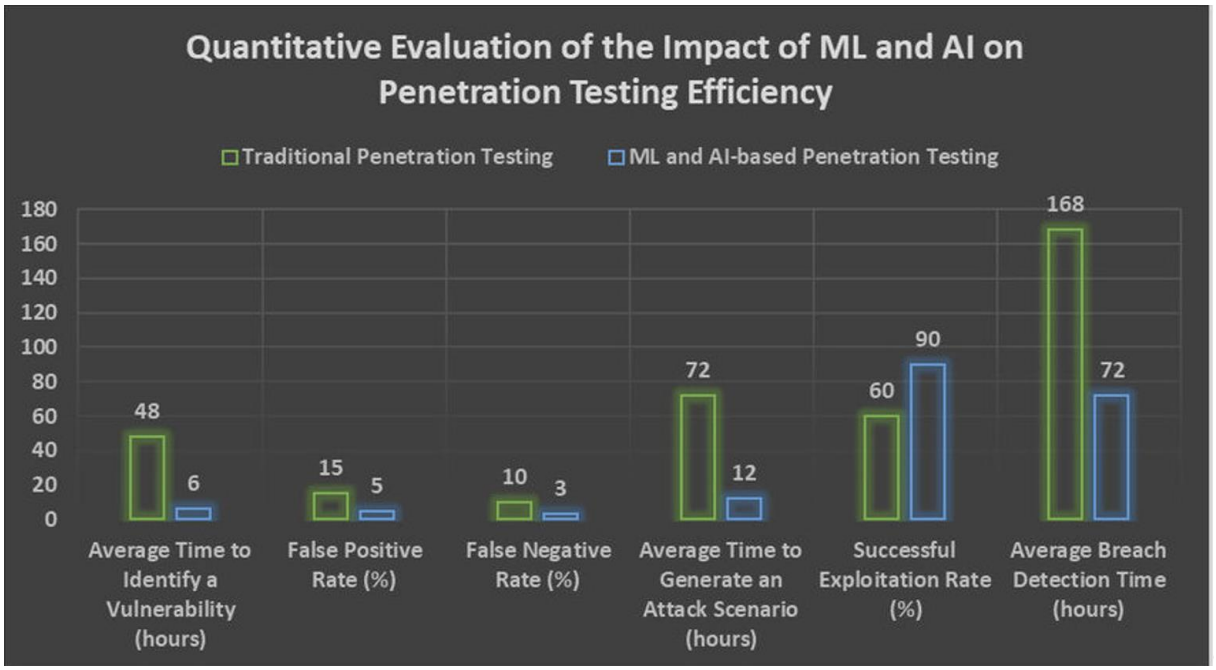

Рис. 1. Кількісна оцінка впливу ML та AI на ефективність тестування на проникнення

Використання ML і AI у тестуванні на проникнення значно підвищує ефективність і точність процесу, скорочуючи час і знижуючи кількість хибнопозитивних та хибнонегативних спрацювань. Це підтверджує переваги впровадження сучасних технологій у кібербезпеку.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Використання методів машинного навчання та великих мовних моделей у процесі тестування на проникнення демонструє значний потенціал для підвищення ефективності, точності та швидкості виявлення вразливостей. Проведений аналіз показав, що такі підходи дозволяють скоротити час ідентифікації загроз, зменшити кількість хибнопозитивних та хибнонегативних спрацювань, а також автоматизувати рутинні завдання, зокрема аналіз логів, генерацію експлойтів та створення звітів. Інтеграція машинного навчання з традиційними методами пентесту надає більш гнучкі та масштабовані рішення для кібербезпеки, сприяючи зосередженню уваги фахівців на критичних аспектах. Однак, незважаючи на досягнутий прогрес, залишаються виклики, які потребують подальшого вирішення. Основними з них є залежність ефективності моделей від якості навчальних даних, висока вартість обчислень, а також необхідність вдосконалення механізмів адаптації

до різних середовищ і типів атак. Особливу увагу слід приділити створенню високоякісних навчальних наборів даних, інтеграції ML та LLM з існуючими CI/CD процесами, а також забезпеченню прозорості та інтерпретованості результатів. Перспективи подальших досліджень включають:

1. Розробку алгоритмів, які зможуть ефективно працювати з обмеженими ресурсами та в умовах динамічного середовища.
2. Поглиблене дослідження нейросимвольних підходів для інтеграції машинного навчання зі статичним та динамічним аналізом.
3. Удосконалення методів моделювання сценаріїв атак із використанням Reinforcement Learning.
4. Розширення можливостей автоматизації виявлення складних вразливостей, що вимагають контекстуального аналізу.
5. Створення спеціалізованих навчальних наборів даних для конкретних типів додатків і середовищ.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Li, Z., Dutta, S., & Naik, M. (2024). *LLM-Assisted static analysis for detecting security vulnerabilities.* https://doi.org/10.48550/arXiv.2405.17238
2. Saini, J., & Bansal, A. (2024). Automated penetration testing: Machine learning approach. *Symposium on Computing & Intelligent Systems (SCI), vol. 3682*, 113–125.
3. Omar, M. (2023). *Detecting software vulnerabilities using Language Models*. https://doi.org/10.48550/arXiv.2302.11773
4. Haidur, H. I., Gakhov, S. O., Marchenko, V. V., & Haidur, K. V. (2024). Conceptual model of detection of phishing attacks based on the use of support vector methods. *Modern Information Security*. https://doi.org/10.31673/2409-7292.2024.020003
5. Burova, N., Oprysk, R., Kurii, Y., Lakh, Y., & Susukailo, V. (2024). Machine learning as a key tool for defensive cyber operations: Effectiveness of phishing threat detection. *Journal of Scientific Papers "Social Development and Security", 14(5),* 113–123. https://doi.org/10.33445/sds.2024.14.5.11
6. Orlivska, V. (2024). Prospects for the application of data mining in cybersecurity. *Information technologies and systems in the documentary field*, 140–142.
7. Johnson, A. A., Ott, M. Q., & Dogucu, M. (2022). Naive bayes classification. *Bayes rules!*, 355–372. https://doi.org/10.1201/9780429288340-14
8. Lunhol, O. (2024). Overview of cybersecurity methods and strategies using artificial intelligence. *Electronic Professional Scientific Journal "Cybersecurity: Education, Science, Technique", 1(25),* 379–389. https://doi.org/10.28925/2663-4023.2024.25.379389
9. Xu, K., Yu, J., Hu, Y., & Ai, X. (2019). Security monitoring data fusion method based on ARIMA and LS-SVM. *IOP Conference Series: Earth and Environmental Science*, *252*, 042104. https://doi.org/10.1088/1755-1315/252/4/042104
10. Tolkachova, A., & Posuvailo, M.-M. (2024). Penetration testing using deep reinforcement learning. *Electronic Professional Scientific Journal "Cybersecurity: Education, Science, Technique", 3(23)* 17–30. https://doi.org/10.28925/2663-4023.2024.23.1730
11. Piskozub, A., Zhuravchak, D., & Tolkachova, A. (2023). Researching vulnerabilities in chatbots with llm (Large language model). *Ukrainian Scientific Journal of Information Security*, *29(3),* 111–117. https://doi.org/10.18372/2225-5036.29.18069
12. Machhindra, P. A., Vijay, B. N., Mahendra, B. S., & Rahul, C. A. (2023). Enhancing cyber security through machine learning: A comprehensive analysis. *Conference: 2023 4th International Conference on Computation, Automation and Knowledge Management (ICCAKM)*. https://doi.org/10.1109/ICCAKM58659.2023.10449547

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 3 (27), 2025
ISSN 2663 - 4023

Anastasiia Zhuravchak
Academic degree, Academic title, Position
Lviv Polytechnic National University, Lviv, Ukraine
ORCID ID: 0000-0002-8196-7963
anastasiia.y.tolkachova@lpnu.ua

Andriian Piskozub
Academic degree, Academic title, Position
Lviv Polytechnic National University, Lviv, Ukraine
ORCID ID: 0000-0002-3582-2835
andriian.z.piskozub@lpnu.ua

ANALYSIS OF MACHINE LEARNING METHODS FOR AUTOMATING PENETRATION TESTING

Abstract. Automation of penetration testing using machine learning methods is one of the most promising areas in modern cybersecurity. The traditional approach to penetration testing requires significant resources, including financial ones, as well as the involvement of highly qualified specialists capable of conducting a comprehensive assessment of system security. This approach may not always provide sufficient speed in detecting new threats, especially in the face of the ever-increasing complexity of cyberattacks and the large number of vulnerabilities. The introduction of machine learning methods into the pentesting process allows creating flexible, adaptive systems that can not only automate routine tasks but also increase the accuracy and efficiency of vulnerability detection. This article provides an overview of the key machine learning algorithms that can be used to automate penetration testing, including support vector machines, random forest, naive Bayes, decision trees, and reinforcement learning methods. Each of these algorithms offers certain advantages in the context of vulnerability analysis, threat classification, and prioritisation of critical security issues. Special attention is paid to the role of large language models in the automation process. They can analyse logs, classify threats, generate reports, and even provide recommendations for fixing identified vulnerabilities. Such models can significantly increase the productivity of specialists by performing routine tasks automatically, which is especially useful when integrated with CI/CD processes. At the same time, the use of LLM has certain limitations, such as dependence on up-to-date data and high computing costs. The article also discusses the challenges and limitations of implementing machine learning algorithms in the pentesting process, such as the need for a large amount of high-quality data to train models, high computing resources, and the risks associated with possible false positives. The results of the study demonstrate that machine learning algorithms have significant potential to improve the efficiency of automated penetration testing, especially in large infrastructures with numerous vulnerabilities.

Keywords: machine learning; penetration testing; deep learning; big language models; decision trees; SVMs.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Li, Z., Dutta, S., & Naik, M. (2024). *LLM-Assisted static analysis for detecting security vulnerabilities.* https://doi.org/10.48550/arXiv.2405.17238
2. Saini, J., & Bansal, A. (2024). Automated penetration testing: Machine learning approach. *Symposium on Computing & Intelligent Systems (SCI), vol. 3682*, 113–125.
3. Omar, M. (2023). *Detecting software vulnerabilities using Language Models*. https://doi.org/10.48550/arXiv.2302.11773
4. Haidur, H. I., Gakhov, S. O., Marchenko, V. V., & Haidur, K. V. (2024). Conceptual model of detection of phishing attacks based on the use of support vector methods. *Modern Information Security*. https://doi.org/10.31673/2409-7292.2024.020003

5. Burova, N., Oprysk, R., Kurii, Y., Lakh, Y., & Susukailo, V. (2024). Machine learning as a key tool for defensive cyber operations: Effectiveness of phishing threat detection. *Journal of Scientific Papers "Social Development and Security", 14(5),* 113–123. https://doi.org/10.33445/sds.2024.14.5.11
6. Orlivska, V. (2024). Prospects for the application of data mining in cybersecurity. *Information technologies and systems in the documentary field,* 140–142.
7. Johnson, A. A., Ott, M. Q., & Dogucu, M. (2022). Naive bayes classification. *Bayes rules!*, 355–372. https://doi.org/10.1201/9780429288340-14
8. Lunhol, O. (2024). Overview of cybersecurity methods and strategies using artificial intelligence. *Electronic Professional Scientific Journal "Cybersecurity: Education, Science, Technique", 1(25),* 379–389. https://doi.org/10.28925/2663-4023.2024.25.379389
9. Xu, K., Yu, J., Hu, Y., & Ai, X. (2019). Security monitoring data fusion method based on ARIMA and LS-SVM. *IOP Conference Series: Earth and Environmental Science*, *252,* 042104. https://doi.org/10.1088/1755-1315/252/4/042104
10. Tolkachova, A., & Posuvailo, M.-M. (2024). Penetration testing using deep reinforcement learning. *Electronic Professional Scientific Journal "Cybersecurity: Education, Science, Technique", 3(23)* 17–30. https://doi.org/10.28925/2663-4023.2024.23.1730
11. Piskozub, A., Zhuravchak, D., & Tolkachova, A. (2023). Researching vulnerabilities in chatbots with llm (Large language model). *Ukrainian Scientific Journal of Information Security*, *29(3),* 111–117. https://doi.org/10.18372/2225-5036.29.18069
12. Machhindra, P. A., Vijay, B. N., Mahendra, B. S., & Rahul, C. A. (2023). Enhancing cyber security through machine learning: A comprehensive analysis. *Conference: 2023 4th International Conference on Computation, Automation and Knowledge Management (ICCAKM)*. https://doi.org/10.1109/ICCAKM58659.2023.10449547