Державний вищий навчальний заклад

«Донецький національний технічний університет»

«Наукові досягнення та відкриття сучасної молоді»

V Всеукраїнська наукова конференція студентів та молодих вчених

Присвячена 105 річниці Державного вищого навчального закладу

«Донецький національний технічний університет»

Збірник матеріалів

12 травня 2026 року

м. Дрогобич

УДК 001-053.6
Н 34

Наукові досягнення та відкриття сучасної молоді [Електронний ресурс] : зб.матер. IV Всеукр. наук. конф. студ. та молодих вчених (Дрогобич, 12 трав. 2026 р.) / Держ. вищ. навч. заклад «Донецький національний технічний університет». – Дрогобич : ДВНЗ «ДонНТУ», 2026. – 172 с.

ISBN: 978-966-377-260-8

Збірник містить доповіді учасників V Всеукраїнської науково-практичної конференції «Наукові досягнення та відкриття сучасної молоді», які розподілені за секціями: «Комп'ютерні та технічні науки», «Соціо-гуманітарні науки, освіта, педагогіка, бізнес, адміністрування та право», «Цивільна та екологічна безпека».

Видання може бути корисним здобувачам вищої освіти, молодим науковцям та викладачам. Усі матеріали друкуються в авторській редакції і відповідальність за їх зміст несуть автори. Оргкомітет конференції претензії з цього приводу не приймає.

Відповідальна за випуск:
Микита АЛЕКСАНДРОВ – Голова Ради молодих вчених ДВНЗ «ДонНТУ», Ph.D, доцент кафедри штучного інтелекту і кібербезпеки.

Організаційний комітет

Попова Ольга Юріївна – д-р екон. наук, професор кафедри управління і фінансово-економічної безпеки, проректор з наукової роботи, **голова оргкомітету**.

Александров Микита Олександрович – Ph.D., доцент кафедри штучного інтелекту і кібербезпеки, **заступник голови оргкомітету.**

Дорогий Ярослав Юрійович – д-р тех. наук, професор, академік МКА, завідувач кафедри штучного інтелекту і кібербезпеки.

Кутняшенко Олексій Ігорович – к.т.н., доцент, доцент кафедри природоохоронної діяльності.

Ігнаткін Валерій Устинович – д.т.н. професор, професор кафедри автоматики і телекомунікацій.

Жуковська Дар'я Олександрівна – старший викладач кафедри автоматики та телекомунікацій.

Єжова Єлизавета Олексіївна – асистент кафедри штучного інтелекту і кібербезпеки.

Хома Дмитро Юрійович – асистент кафедри штучного інтелекту і кібербезпеки.

 © ДВНЗ «ДонНТУ», 2026

ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ЗАДАЧ ПЕНТЕСТИНГУ

[1]Кравчук Владислав Сергійович, аспірант
[2]Дорогий Ярослав Юрійович, д.т.н., проф., зав. каф. ШІК
ДВНЗ «Донецький національний технічний університет»
[1]vladyslav.kravchuk.asp@donntu.edu.ua, [2]yaroslav.dorohyi@donntu.edu.ua

Анотація. Об'єктом дослідження є застосування ШІ для автоматизації задач тестування на проникнення. Метою роботи є аналіз сучасних підходів до використання машинного навчання (ML), великих мовних моделей (LLM) та підкріплювального навчання (RL) у розвідці, скануванні, експлуатації, післяексплуатації та звітуванні. Методи дослідження охоплюють критичний огляд літератури, класифікацію інструментів за етапами пентестингу та порівняльний аналіз їх ефективності. Ключові результати демонструють значне прискорення процесу пентестингу завдяки LLM (наприклад, PentestGPT) та RL-агентам, однак виявлено прогалини в інтеграції технологій та оцінці ризиків.

Ключові слова: пентестинг, штучний інтелект, тестування на проникнення, вразливості, автоматизація, кібербезпека, машинне навчання, великі мовні моделі, підкріплювальне навчання.

Постановка проблеми дослідження. Сучасний ландшафт кіберзагроз характеризується швидким зростанням кількості та складності атак: за даними глобальних звітів, щорічно фіксується понад 2,5 млрд інцидентів, пов'язаних із вразливостями програмного забезпечення та інфраструктури. Ручний пентестинг, традиційно покладений на кваліфікованих фахівців, не встигає за масштабом і швидкістю поширення загроз через обмежену швидкість виконання, високу трудомісткість та суб'єктивність оцінки.

Сучасні атаки, зокрема APT та zero-day експлойти, вимагають оперативної адаптації стратегій, що суперечить браку кваліфікованих спеціалістів у галузі кібербезпеки: за оцінками, дефіцит кадрів сягає мільйонів фахівців глобально. Ці протиріччя актуалізують потребу в автоматизації пентестингу за допомогою ШІ, яка здатна поєднувати швидкість обробки даних з імітацією людського міркування.

Аналіз літературних джерел та публікацій. Аналіз літератури 2021–2025 рр. свідчить про активне впровадження ML/LLM/RL у пентестинг. Автори [1] запропонували PentestGPT – LLM-фреймворк, що автоматизує всі етапи пентестингу шляхом ітеративного міркування, демонструючи зростання ефективності на 228,6 % порівняно з базовими моделями. В роботі [2] емпірично оцінено можливості LLM у пентестингу, виявивши їхню спроможність генерувати реалістичні команди, але зазначили проблеми з галюцинаціями та відсутністю контекстуальної пам'яті. У сфері RL автори [3] розробили ієрархічний RL-фреймворк для зменшення простору дій і стану, що дозволило ефективніше знаходити оптимальні шляхи атаки в симульованих мережах. В праці [4] представлено фреймворк IAPTF) Intelligent Automated Penetration Testing Framework) на базі побудованій на моделі RL, який автоматизує

послідовне прийняття рішень у великих мережах. Дослідники в [5] поєднали RL з LLM для автономного пентестингу, продемонструвавши переваги гібридного підходу в багаторівневих системах. В праці [6] застосувано глибокий RL для автоматизації розвідки та експлуатації, досягнувши точності 0,86 у виборі оптимальних шляхів. Автори [7] запропонували ієрархічний глибокий RL з експертними знаннями, що підвищило ефективність у складних середовищах. Дослідники [8] у систематичному огляді RL-підходів до автономного пентестингу підкреслили прогрес у моделюванні середовища, але вказали на обмеженість реальних тестів. Автор [9] розглянув генеративний ШІ у SDLC, включаючи побудоване ШІ тестування на проникнення для малого бізнесу. В праці [10] описано таксономію генеративного ШІ у для оцінки вразливостей, виокремлено атакуючий та захисний застосування LLM.

Критичний огляд розглянутих джерел виявив деякі прогалини: більшість досліджень зосереджені на окремих етапах (переважно exploitation), бракує комплексних фреймворків, що охоплюють весь цикл пентестингу; оцінка проводиться переважно в симульованих середовищах, а не на реальних цілях через етичні обмеження; недостатньо вивчено стійкість ШІ-інструментів до змагальних атак та етичні/правові наслідки автономної експлуатації; рідко аналізуються ризики галюцинацій LLM у критичних задачах звітності.

Мета та задачі дослідження. Метою дослідження є класифікація та оцінка ефективності інструментів штучного інтелекту для автоматизації задач пентестингу в умовах сучасних кіберзагроз. Для досягнення мети поставлено такі задачі: проаналізувати сучасні наукові публікації щодо застосування ML/LLM/RL у пентестингу; класифікувати інструменти ШІ за етапами пентестингу; порівняти ключові інструменти за критеріями ефективності, охоплення та ризиків; оцінити практичну ефективність та потенційні загрози використання ШІ; запропонувати напрями інтеграції гібридних підходів для red team та blue team.

Результати дослідження. Результати дослідження дозволяють класифікувати інструменти ШІ за стандартними етапами пентестингу (PTES): reconnaissance (розвідка), scanning (сканування), exploitation (експлуатація), post-exploitation (післяексплуатація) та reporting (звітування). На етапі розвідки LLM (наприклад, PentestGPT) автоматизують OSINT та генерацію цілей. Сканування підтримується RL-агентами, які оптимізують вибір інструментів (nmap, Nessus). Експлуатація є найрозвинутішою сферою завдяки ієрархічному RL, що моделює послідовність дій у графі вразливостей. Післяексплуатація включає LLM для закріплення та латерального руху. Звітування автоматизується генеративними моделями для створення структурованих звітів.

Порівняльна характеристика ключових інструментів наведена в таблиці 1.

Таблиця 1 – Ключові інструменти для етапів пентестингу

Інструмент	Етапи пентестингу	Технологія	Ефективність (за літературою)	Ризики
PentestGPT	Всі етапи	LLM (GPT-4)	+228,6 % виконання завдань	Галюцинації, недетермінованість
Hierarchical RL Framework (Liu et al.)	Recon, Exploitation, Post-exploitation	Hierarchical RL	Зменшення простору дій, висока точність	Потреба в навчанні, overfitting
IAPTF (Ghanem)	Scanning, Exploitation	Model-based RL	Ефективний для великих мереж	Обмежена адаптивність до zero-day
Deep RL (Nhu et al.)	Recon, Exploitation	Deep Q-Network	Точність 0,86 у виборі шляхів	Високі обчислювальні витрати
AutoPentester (2025)	Всі етапи (агентний)	LLM-агент	Перевищує PentestGPT на 19,8 % (за відгуками)	Залежність від API, етичні ризики

Оцінка ефективності показує, що LLM-фреймворки значно прискорюють процес (до кількох годин замість днів), проте RL забезпечує кращу адаптивність у динамічних середовищах. Основні ризики – технічні (галюцинації, помилкові експлойти), операційні (брак пояснюваності) та правові (автономна дія без явної авторизації).

Висновки. Дослідження доводить, що інструменти ШІ здатні автоматизувати значну частину задач пентестингу, забезпечуючи вищу швидкість і масштабованість порівняно з ручними методами. Практична цінність для red team полягає в прискоренні розвідки та експлуатації, а для blue team – у симуляції реалістичних атак для покращення оборонних стратегій. Водночас використання ШІ у пентестингу підлягає суворим етичним та правовим обмеженням: автономні дії можливі лише за наявності письмової авторизації, з дотриманням норм GDPR, Закону України «Про основні засади забезпечення кібербезпеки» та міжнародних стандартів (ISO 27001); неконтрольоване застосування може призвести до порушення законодавства та етичних норм. Напрямами подальших досліджень є створення гібридних LLM+RL-фреймворків з вбудованою пояснюваністю, розробка бенчмарків на реальних цілях, а також вивчення захисту самих ШІ-інструментів від змагальних атак.

ЛІТЕРАТУРА

1. PentestGPT: Evaluating and Harnessing Large Language Models for Automated Penetration Testing / G. Deng et al. *arXiv*. 2024. DOI: 10.48550/arXiv.2308.06782.

2. Happe A., Cito J. Getting pwn'd by AI: Penetration Testing with Large Language Models. *Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2023)* (New York, NY, USA, 11–15 December 2023). New York : ACM, 2023. P. 2082–2086. DOI: 10.1145/3611643.3613083.

3. An Automated Penetration Testing Framework Based on Hierarchical Reinforcement Learning / H. Liu et al. *Electronics*. 2024. Vol. 13, No. 21. Art. 4311. DOI: 10.3390/electronics13214311.

4. Ghanem M. C. Hierarchical reinforcement learning for efficient and effective automated penetration testing of large networks. *Journal of Intelligent Information Systems*. 2023. Vol. 60, No. 2. P. 281–303. DOI: 10.1007/s10844-022-00738-0.

5. Analysis of Autonomous Penetration Testing Through Reinforcement Learning and Large Language Models / A. C. Moreno et al. *Sensors*. 2025. Vol. 25, No. 1. Art. 211. DOI: 10.3390/s25010211.

6. Leveraging Deep Reinforcement Learning for Automating Penetration Testing in Reconnaissance and Exploitation Phase / L. V. Hoang et al. *2022 RIVF International Conference on Computing and Communication Technologies (RIVF)* (Ho Chi Minh City, Vietnam, 7–9 November 2022). IEEE, 2022. P. 41–46. DOI: 10.1109/RIVF55975.2022.10013801.

7. A Hierarchical Deep Reinforcement Learning Model with Expert Prior Knowledge for Intelligent Penetration Testing / Q. Li et al. *Computers & Security*. 2023. 103358. DOI: 10.1016/j.cose.2023.103358.

8. Autonomous penetration testing using reinforcement learning: A review and perspectives / J. Liu et al. *Expert Systems with Applications*. 2025. Vol. 300. Art. 130219. DOI: 10.1016/j.eswa.2025.130219.

9. Exploring the role of generative AI in enhancing cybersecurity in software development life cycle / H. A. Al-Hashimi et al. *Array*. 2025. Vol. 28. 100509. DOI: 10.1016/j.array.2025.100509.

10. Cybersecurity in the age of generative AI: A systematic taxonomy of AI-powered vulnerability assessment and risk management / S. L. Mirtaheri et al. *Future Generation Computer Systems*. 2026. Vol. 175. 108107. DOI: 10.1016/j.future.2025.108107.