

The Use of Artificial Intelligence Techniques for Detecting Web Application Security Vulnerabilities and Improving Penetration Testing Efficiency: An Applied Study

Literature Review

1. Sanchez Collado et al. (2024)

Sanchez Collado et al. (2024) performed a systematic review on the use of AI techniques in pentesting web applications from 2013 to 2024. The study summarized the results from the previous studies and analyzed the effectiveness of machine learning and other artificial intelligence techniques in improving the penetration testing processes versus the traditional manual methods. The review revealed that AI has an impact on improving testing efficiency, automating the discovery of vulnerabilities, saving test time, and enhancing the precision of identifying potential vulnerabilities. Additionally, AI-powered penetration testing helps organizations enhance web application security without putting excessive strain on human resources. However, the authors found that there were also some obstacles, such as the lack of model generalization, data quality problems, lack of explainability, and challenges integrating AI into real-world security environments. The review concluded that while AI is a promising direction in intelligent penetration testing, it still needs to be developed to be more robust, transparent and usable in the enterprise cybersecurity landscape.

2. Saber, Bahaa-Eldin, ElSayad, & Fayed (2025)

In Saber et al. (2025), the authors examine the effectiveness of an intelligent Nmap-based vulnerability scanning framework that they have developed in comparison with the commercial Nessus vulnerability scanner. The proposed framework was designed to enhance the performance of vulnerability detection by adopting automated scanning techniques and intelligent matching algorithms. Results of the experiments showed that the framework successfully finished the fully completed vulnerability assessments in about four minutes, while, the Nessus took nearly sixty-seven minutes. Although the proposed framework reduced execution time considerably, it still retained a high detection accuracy rate of 95% with 37 of 39 vulnerabilities detected. The results suggest that intelligent automation can be a gamechanger in expediting vulnerability assessment at a level that won't compromise its accuracy. This study identifies the potential of integrating AI-driven scan mechanisms into penetration testing to drive gains in cybersecurity defenses and shorten the time required to assess cybersecurity.

3. Malkawi & Alhajj (2026)

This Systematic Review aims to analyze the techniques, challenges, and future trends related to AI-powered Vulnerability Detection and Patch Management in Cybersecurity. The objective of this Systematic Review is to examine the techniques, challenges, and emerging trends in AI-powered Vulnerability Detection and Patch Management in Cybersecurity.

Malkawi and Alhadj (2026) conducted a thorough systematic review of vulnerability detection and automated patching methods for AI-based solutions, examining 29 peer-reviewed papers from 2019 to 2024. The review focused on vulnerability identification and remediation using machine learning, deep learning, reinforcement learning, neuro-symbolic AI, and large language models. The results showed that the accuracy, scalability, and response time of AI-based systems are significantly better than traditional rule-based security systems. The authors also touched on automated vulnerability prioritization and intelligent patch generation using AI. There are however important limitations, such as a lack of benchmark datasets, poor generalizability of models, lack of adversarial robustness, limited transparency and explainability. The study highlighted the need for building trustworthy, explainable, and robust AI systems that can aid in real-world vulnerability management in various cybersecurity scenarios.

4. Radanliev (2025)

Radanliev (2025) has suggested an evolving collaborative penetration testing scheme for the secure development of emerging generative artificial intelligence systems against advanced cyber threats and future quantum computing attacks. It incorporates several security assessment techniques such as static application security testing (SAST), dynamic application security testing (DAST), interactive application security testing (IAST), blockchain logging, and AI-powered red team simulations. The effectiveness of the framework was proved by experimental evaluation, which identified and addressed over 300 security flaws, and eliminated about 70% of high severity flaws within two weeks of monitoring using blockchain. Further, the quantum resistant cryptographic methods effectively thwarted simulated quantum attacks to applications. The study clearly showed the value of combining AI with robust penetration testing processes in identifying vulnerabilities, mitigating them, conducting security audits, and boosting cyber resilience.

5. Agale, Kadam, Poyeraker & Deshpande (2026).

Agale et al. (2026) discussed the recent progress of AI-assisted and automated penetration testing tools, highlighting the intelligent security assessment frameworks based on machine learning, large language models, and automated security tools. The review introduced the concept of smart systems that automate reconnaissance, vulnerability scanning, exploitation planning and contextual reasoning, which replace manual methods of penetration testing. The authors analyzed and contrasted the state-of-the-art AI-driven penetration testing frameworks in terms of features, scalability, real-world use, and enterprise suitability. The study highlights recurring issues that include apparent weaknesses in the explainability, ethical concerns with the models and the complexity of integration, as well as having false positives when it comes to finding vulnerabilities. While AI and LLM can be very effective in improving the efficiency of penetration

testing and decision-making, there is still a need for more research to ensure reliability, transparency, scalability and operational adoption in the current cyber security landscape.

Summary of Previous Studies

Author(s) & Year	Objective	Methodology	Key Findings	Limitation
Sanchez Collado et al. (2024)	Review AI applications in web application penetration testing.	Systematic literature review.	AI improves detection accuracy, automation, and testing efficiency.	No practical implementation or experimental validation.
Saber et al. (2025)	Compare an AI-assisted Nmap framework with Nessus.	Experimental comparative study.	Reduced scan time while maintaining 95% detection accuracy.	Focused only on vulnerability scanning.
Malkawi & Alhajj (2026)	Review AI techniques for vulnerability detection and patch management.	Systematic literature review of 29 studies.	AI enhances precision, scalability, and automated remediation.	Limited real-world validation and benchmark availability.
Radanliev (2025)	Develop an AI-based collaborative penetration testing framework.	Applied experimental framework.	Detected over 300 vulnerabilities and improved remediation efficiency.	Focused mainly on generative AI and quantum security.
Agale et al. (2026)	Review AI-assisted and automated penetration testing techniques.	Literature review.	LLMs and AI improve automation and decision-making.	No empirical evaluation; discusses challenges only.

Comparative Analysis

Artificial Intelligence (AI) has been widely proven to enhance the efficiency of vulnerability detection and penetration testing. Sanchez Collado et al. (2024) and Malkawi and Alhajj (2026) have performed systematic reviews demonstrating the benefits of using AI in improving detection accuracy and the automation of the process, but also identified issues with explainability and practical implementation. Although Saber et al. (2025) experimented with vulnerability scanning alone, they found that an AI-assisted Nmap framework decreases the scanning time significantly while retaining high accuracy of detection. Radanliev (2025) also introduced him an integrated AI-driven penetration testing framework that integrates multiple security testing techniques, and Agale et al. (2026) explored approaches to penetration testing using AI and language models (LLMs), highlighting the key trends and limitations in the field. To sum up, the previous studies validate AI's promise and are limited in scope of providing comprehensive applied evaluations for web application penetration tests.

Research Gap

While AI's potential in cybersecurity is evident from prior studies, there are still some areas to be filled. Most studies are systematic reviews and not applied research, and experimental studies may concentrate only on vulnerability scanning and not the entire penetration testing process. Furthermore, there are limited studies that have compared various AI techniques in real-world environments of a web application, and issues of explainability, scalability and generalizability are not addressed. To fill these gaps, this study aims to perform an applied assessment of AI techniques for detecting web application security vulnerabilities and increasing the efficiency of penetration testing, and offer practical examples of their effectiveness over conventional methods.

References

Agale, A., Kadam, A., Poyeraker, V., & Deshpande, R. (2026). *Review of AI-assisted and automated penetration testing techniques. International Journal of Advanced Research in Science, Communication and Technology*, 6(4).

Malkawi, M., & Alhajj, R. (2026). AI-powered vulnerability detection and patch management in cybersecurity: A systematic review of techniques, challenges, and emerging trends. *Machine Learning and Knowledge Extraction*, 8(1), Article 19.

Radanliev, P. (2025). Collaborative penetration testing suite for emerging generative AI algorithms. *Applied Intelligence*, 55, Article 1030.

Saber, V., Bahaa-Eldin, A. M., ElSayad, D., & Fayed, Z. T. (2025). Exploring the correlation between vulnerability scanning and Nmap. *International Journal of Intelligent Computing and Information Sciences*, 25(1), 41–50.

Sanchez Collado, G., Olayinka, O., & Pasikhani, A. (2024). Web application penetration testing with artificial intelligence: A systematic review. In *Proceedings of the 22nd International Symposium on Network Computing and Applications (NCA 2024)*. IEEE.