AI-Enhanced Web Vulnerability Scanner: Integrating Dynamic Payload Generation and Machine Learning-Based Risk Assessment

Shreyas Malusare[1], Hrushikesh Kolhe[2], Priyanka Belapurkar[3], Prof. Ulka Bansode[4]
[1,2,3,4]*Computer Engineering, KJ College of Engineering and Management Research, Pune, India*

Abstract*—Modern web ecosystems confront an unprecedented surge in sophisticated cyber threats, predominantly characterized by multi-vector injection flaws, cross-site scripting, and security misconfigurations. While automated Dynamic Application Security Testing (DAST) utilities are routinely deployed to audit web surfaces, conventional black-box scanners suffer from significant operational bottlenecks. These tools rely heavily on static payload dictionaries, yield substantial false-positive rates, and possess a fundamental blindness to business context—reporting technical flaw severity via static scores while failing to evaluate the operational impact on the targeted asset. To bridge this critical gap, this research introduces an intelligent, web-based automated vulnerability scanner. The proposed architecture integrates an asynchronous reconnaissance engine with an environment-aware fuzzer, replacing generic wordlists with context-prioritized payload execution. Furthermore, an advanced machine learning classification module evaluates the contextual architecture of discovered vulnerable endpoints to dynamically predict the asset's business value and magnitude of impact.

***Index Terms*—Dynamic Application Security Testing, Machine Learning, Risk Assessment, Vulnerability Scanning, Web Security.**

I. INTRODUCTION

The global digital landscape relies extensively on complex, multi-tier web applications to power critical socioeconomic sectors. This continuous expansion of web surfaces has concurrently amplified the available cyber-attack envelope. Injection anomalies—encompassing Structured Query Language Injection (SQLi) and Cross-Site Scripting (XSS)—persistently rank as severe systemic hazards within the Open Web Application Security Project (OWASP) top threat matrices. To mitigate these security flaws prior to production deployment, organizations implement automated Web Vulnerability Scanners (WVSs) operating on black-box, Dynamic Application Security Testing (DAST) principles. Black-box testing probes functional software environments externally without direct access to backend source code repositories. While highly practical, comprehensive academic evaluations reveal severe architectural limitations in existing DAST platforms. Conventional scanners execute static dictionary attacks, indiscriminately parsing dense payload lists against every discovered input vector without structural contextual awareness. Consequently, these systems incur massive operational time overheads and generate an overwhelming volume of false-positive indications. More critically, contemporary scanners suffer from acute business context blindness. Standard tools assign invariant Common Vulnerability Scoring System (CVSS) numbers based entirely on technical exploitability variables. However, they cannot automatically deduce whether a vulnerable endpoint interfaces with low-risk generic content or handles high-value sensitive data. Determining actual business risk dictates an intricate evaluation of both threat likelihood and the target asset's distinct data sensitivity. This research directly addresses these long-standing security engineering limitations by introducing a novel, AI-Enhanced Web Vulnerability Scanner. The proposed system replaces primitive brute-force fuzzing with a context-aware injection engine and integrates a machine learning module to automate asset business value classification.

By combining technical severity data with predictive impact parameters, the scanner generates dynamic, prioritized risk metrics aligned closely with

standardized risk frameworks. The system introduces an intelligent workflow where discovered endpoints are analyzed, security testing is performed using optimized payload strategies, and detected vulnerabilities are evaluated using an automated risk prioritization mechanism. This helps security professionals focus on high-impact vulnerabilities first and improves the efficiency of the overall security assessment process.

II. MOTIVATION AND PROBLEM STATEMENT

The rapid digital transformation across global enterprise, healthcare, and governance sectors has shifted critical infrastructure entirely onto complex, multi-tier web applications. Modern web architectures, heavily reliant on integrated APIs and dynamic frontend frameworks, are deployed through Continuous Integration/Continuous Deployment (CI/CD) pipelines. This accelerated development lifecycle has inadvertently expanded the cyber-attack surface, leaving applications susceptible to severe vulnerabilities such as injection flaws, cross-site scripting, and broken access controls.

While manual penetration testing provides deep security insights, it is fundamentally unscalable and cannot match the rapid pace of modern software releases. Consequently, the cybersecurity industry relies heavily on automated Dynamic Application Security Testing (DAST) utilities to audit web surfaces. However, as threat vectors become increasingly sophisticated, the traditional mechanisms powering these automated scanners have proven inadequate. The motivation for this research stems from the critical industry need to evolve DAST tools from rigid, brute-force execution scripts into intelligent, context-aware security platforms. There is a pressing necessity to reduce the immense manual triage burden placed on security engineering teams by providing automated, prioritized, and business-aligned threat intelligence rather than unstructured vulnerability logs.

Despite widespread adoption, current state-of-the-art black-box vulnerability scanners exhibit fundamental architectural and operational deficiencies that severely degrade their utility in agile development environments.

III. LITERATURE REVIEW

Recent engineering advancements in IEEE literature between 2023 and 2026 demonstrate a significant paradigm shift in Dynamic Application Security Testing (DAST). The cybersecurity industry is rapidly transitioning from static, deterministic scanning methodologies toward adaptive, artificial intelligence-driven vulnerability assessment frameworks.

Addressing the limitations of traditional payload injection, Shahrivar et al. [1] engineered a machine-learning-based approach for detecting and mitigating automated DAST attacks. Their research highlighted the fundamentally brittle nature of static, threshold-based scanning engines, which fail to adapt to modern Web 2.0 application behaviors and frequently generate massive computational latency. Expanding on the necessity for intelligent detection automation, Dibouliya et al. [2] proposed a comprehensive machine learning framework specifically designed to detect complex web vulnerabilities, such as SQL Injection (SQLi) and Cross-Site Scripting (XSS). Their empirical evaluations proved that neural network classifications can vastly improve diagnostic accuracy and reduce false-positive rates compared to traditional rule-based heuristics. However, while these studies successfully applied artificial intelligence to the *detection* phase of penetration testing, they largely treated all discovered vulnerabilities with equal operational precedence, failing to address the subsequent challenge of business risk calculation. The critical challenge of context blindness has consequently driven the latest IEEE research toward automated vulnerability prioritization. Bennouk et al. [3] highlighted the inadequacy of utilizing static datasets for threat modeling. Their research demonstrated that employing dynamic data updates and weight optimization significantly improves the predictive accuracy of vulnerability exploitability. Pushing contextual awareness further, Huang et al. [4] introduced AegisGuard, a framework for semantic vulnerability detection and risk stratification that dynamically adjusts reasoning based on system-specific privilege settings. Finally, tackling the problem of overwhelming security alerts, Shimizu and Hashimoto [5] proposed Vulnerability Management Chaining, an integrated decision-tree framework designed for efficient cybersecurity risk prioritization. Their system underscores the necessity of chaining

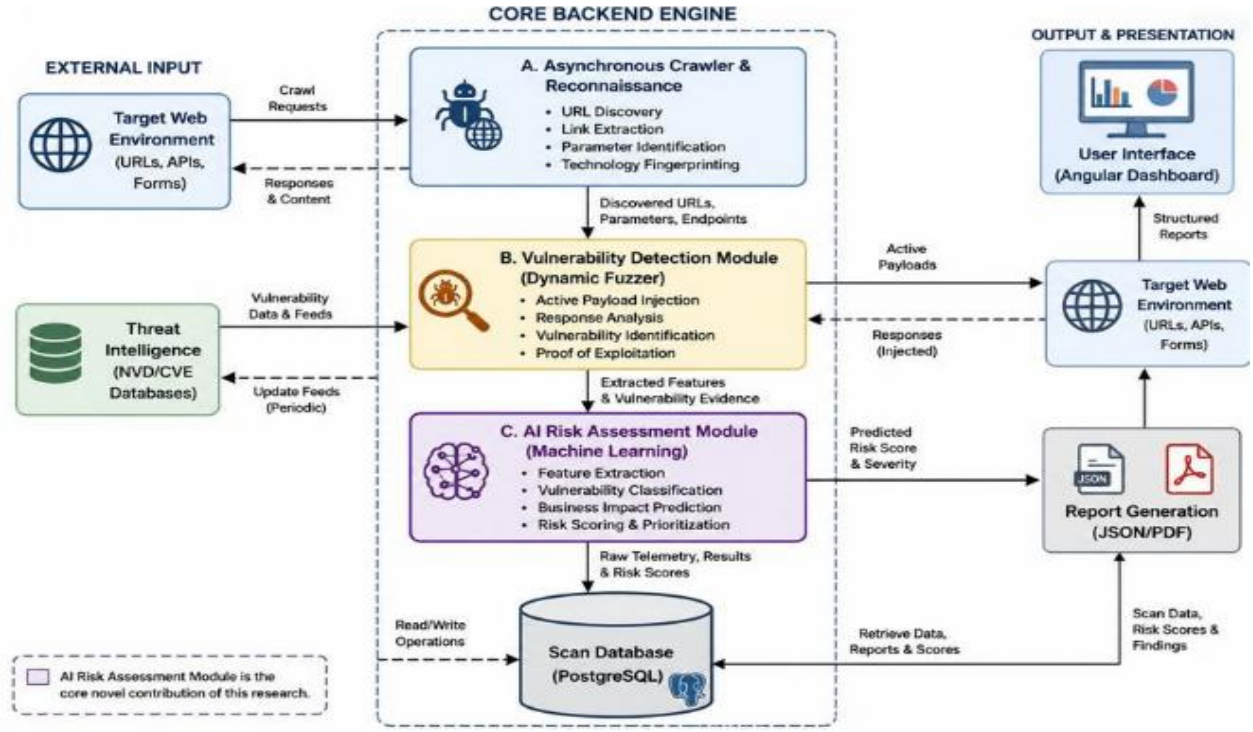
threat intelligence to map detected vulnerabilities to environmental scoring metrics. Shimizu and Hashimoto introduced an integrated decision-tree framework tailored for efficient cybersecurity risk prioritization. Their system underscores the necessity of utilizing machine learning algorithms to autonomously map detected technical vulnerabilities to environmental scoring metrics, such as the Exploit Prediction Scoring System (EPSS) and CVSS.

Despite these advancements, a critical void remains. While machine learning has been successfully applied to either vulnerability detection [1], [2] or threat prioritization [3], [4], [5] independently, there is a conspicuous absence of an integrated, end-to-end automated scanning architecture that simultaneously performs dynamic DAST fuzzing and intelligent business risk prioritization within a single platform. The academic hyper-focus on specific injection flaws, combined with the universal neglect of scanner usability and predictive impact modeling, necessitates a paradigm shift. Modern agile development environments require a departure from brute-force, context-blind scanners towards adaptive, AI-driven security platforms that can autonomously prioritize threats based on realistic operational impact.

In summary, the current state-of-the-art literature reveals a fragmented approach to intelligent security testing. While machine learning has been successfully, but independently, applied to either vulnerability detection [1], [2] or threat prioritization [3], [4], there remains a critical academic and industrial void. There is a conspicuous absence of an integrated, end-to-end automated scanning architecture that simultaneously generates dynamic DAST payloads *and* performs intelligent, machine-learning-driven business risk prioritization within a single, usable platform.

IV. WORKFLOW DIAGRAM DESCRIPTION

The workflow of the proposed AI-Enhanced Web Vulnerability Scanner represents the complete operational pipeline starting from target identification to vulnerability analysis and security report generation. The system follows a modular architecture where each component performs a specific security assessment task and communicates with other modules through structured data exchange. The workflow begins with the Target Web Environment, which acts as the primary input source for the scanning process. The target environment consists of web application URLs, APIs, forms, and accessible endpoints that require security assessment. The system collects this information through the user interface and forwards it to the core scanning engine for further analysis.



The first stage of the scanning process is performed by the Asynchronous Crawler and Reconnaissance Module. This module performs automated discovery of application resources by crawling the target environment. It identifies available URLs, hidden directories, API endpoints, input parameters, forms, and application technologies. The asynchronous approach enables efficient resource discovery by allowing multiple requests to be processed concurrently, reducing the overall scanning time. After reconnaissance, the discovered application structure is transferred to the Vulnerability Detection Module (Dynamic Fuzzer). This module performs active security testing by injecting specially designed payloads into identified parameters and endpoints. The dynamic fuzzing mechanism analyzes server responses, detects abnormal behavior, and identifies potential vulnerabilities such as SQL Injection, Cross-Site Scripting (XSS), insecure configurations, and other web security weaknesses. The detection module also communicates with the Threat Intelligence Database, which contains updated vulnerability information from sources such as National Vulnerability Database (NVD) and Common Vulnerabilities and Exposures (CVE). The detected vulnerability information is then passed to the AI Risk Assessment Module. This module represents the core contribution of the proposed system. Machine learning techniques are applied to analyze extracted vulnerability features, endpoint characteristics, and application context. The model predicts the business impact level of detected vulnerabilities and classifies them according to risk severity. Finally, the processed vulnerability information is forwarded to the Report Generation Module. The generated reports are available in machine-readable and user-friendly formats such as JSON and PDF. The final output is presented through the User Interface (Angular Dashboard), where users can visualize scan results, monitor detected vulnerabilities, and analyze security risks through an interactive dashboard.

V. PROPOSED SYSTEM

The proposed system introduces an AI-Enhanced Web Vulnerability Scanner that integrates automated reconnaissance, dynamic security testing, threat intelligence, and web-based risk assessment into a unified vulnerability analysis framework. The objective of the proposed system is to overcome the limitations of conventional vulnerability scanners by improving detection accuracy.

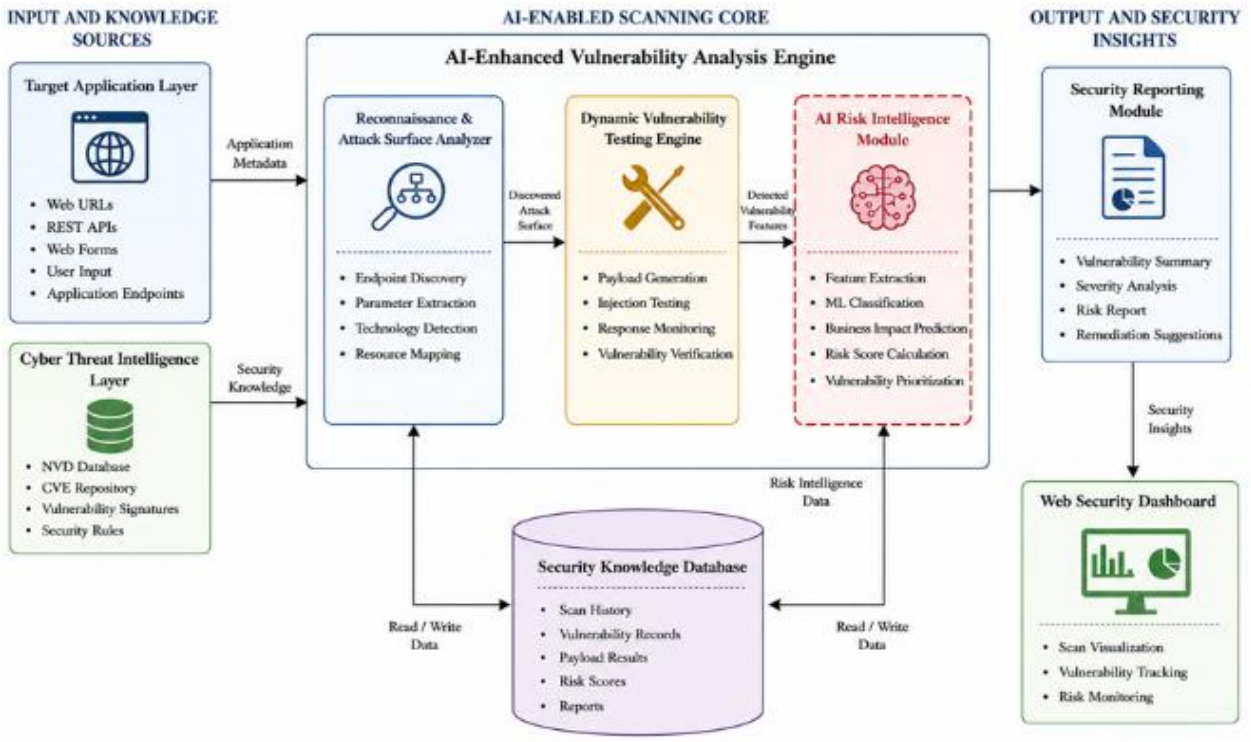


Fig. 2 Proposed System Architecture of AI-Enhanced Web Vulnerability Scanner

VI. OUTCOME

The proposed AI-Enhanced Web Vulnerability Scanner aims to provide an intelligent and automated security assessment framework capable of identifying, analyzing, and prioritizing vulnerabilities present in modern web applications. The system combines traditional vulnerability assessment techniques with artificial intelligence-based analysis to improve the effectiveness of automated security testing.

The primary outcome of the proposed system is the development of an automated vulnerability detection platform that reduces dependency on manual security testing and enables continuous assessment of web applications. By integrating reconnaissance, dynamic vulnerability testing, threat intelligence, and machine learning-based risk analysis, the system provides a complete security evaluation environment.

Automated Attack Surface Discovery-The system performs automated reconnaissance of target web applications by identifying accessible URLs, APIs, forms, parameters, and application technologies. This helps in creating an accurate representation of the application attack surface and provides essential information required for further vulnerability analysis.

A major outcome of the proposed system is the introduction of machine learning-based vulnerability prioritization. Instead of treating all detected vulnerabilities equally, the AI Risk Intelligence Module analyzes vulnerability characteristics and predicts the potential impact level.

VII. METHODOLOGY

Web-Based System Architecture:

The system consists of a frontend dashboard, backend scanning engine, vulnerability analysis modules, database layer, and reporting system. Users interact with the system through the dashboard to submit target applications such as URLs, APIs, and web endpoints. The backend manages scanning operations, security analysis, and communication between different modules.

Target Analysis and Reconnaissance:

The scanning process begins with collecting target application information. The Asynchronous Crawler and Reconnaissance Module analyzes the application structure and identifies resources such as pages, APIs, parameters, forms, and technologies.

Dynamic Vulnerability Detection:

The identified endpoints are processed by the Dynamic Vulnerability Detection Engine. The module performs automated security testing by generating payloads, injecting them into application inputs, and analyzing responses. The system detects vulnerabilities including: SQL Injection, Cross-Site Scripting (XSS), Input Validation Issues, Security Misconfigurations.

Threat Intelligence Integration:

The system integrates security knowledge from sources such as NVD and CVE repositories. This information provides updated vulnerability patterns and security rules, improving detection accuracy and maintaining current threat awareness.

AI-Based Risk Assessment:

The major contribution of the proposed system is the AI Risk Assessment Module. The module extracts vulnerability features such as vulnerability type, affected endpoint, severity indicators, and application context. Machine learning techniques analyze these features to classify vulnerabilities, predict business impact, calculate risk scores, and prioritize critical security issues.

Data Management and Reporting:

The system stores scan history, vulnerability details, payload results, and risk scores in a centralized security database. The processed results are converted into structured reports containing vulnerability information, severity levels, and remediation suggestions.

Web Security Dashboard:

The final results are displayed through the web dashboard, enabling users to monitor application security status and make informed security decisions.

VIII. CONCLUSION

The proposed AI-Enhanced Web Vulnerability Scanner provides an intelligent approach for automated web application security assessment. The

system integrates dynamic payload generation, vulnerability detection, threat intelligence.

ACKNOWLEDGMENT

The authors gratefully acknowledge the guidance of Prof. Ulka Bansode and the Department of Computer Engineering, K. J. College of Engineering and Management Research, Pune, for their invaluable support.

REFERENCES

[1] T. O. Odion, A. I. Ahmed, I. O. Ebo, U. N. Musa, and R. M. Imam, "VulScan: A Web-Based Vulnerability Multi-Scanner for Web Application," in *Proc. Int. Conf. Science, Engineering and Business for Sustainable Development Goals (SEB-SDG)*, Paris, France, 2023, pp. 1–6.

[2] P. Shahrivar, S. Millar, and E. Shereen, "Detecting Web Application DAST Attacks with Machine Learning," in Proc. IEEE Conf. Dependable and Secure Computing (DSC), Nov. 2023.

[3] A. Dibouliya, V. Jotwani, and A. K. Gupta, "Machine Learning for Web Vulnerability Detection," in Proc. Int. Conf. Emerging Technologies and Innovation for Sustainability (EmergIN), 2025.

[4] K. Bennouk, D. Mahouachi, N. Ait Aali, et al., "Dynamic Data Updates and Weight Optimization for Predicting Vulnerability Exploitability," IEEE Access, vol. 13, pp. 65266–65284, 2025.

[5] N. Shimizu and M. Hashimoto, "Vulnerability Management Chaining: An Integrated Framework for Efficient Cybersecurity Risk Prioritization," *IEEE Access*, vol. 14, pp. 31407–31424, Mar. 2026.