International Journal of Scientific Research and Engineering Development-– Volume 9 Issue 6, Nov – Dec 2026
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Artificial Intelligence in Cybersecurity: Techniques, Threat Detection Applications and Comparison with Traditional Approaches

Mopurappagari Bhargav Reddy
Bachelor of Computer Applications
Padmashree Institute of Management & Sciences, Bengaluru
m.bhargavreddy02@gmail.com

Tilak N
Bachelor of Computer Applications
Padmashree Institute of Management & Sciences, Bengaluru
tillu3158@gmail.com

Aman Mahto
Bachelor of Computer Applications
Padmashree Institute of Management & Sciences, Bengaluru
aman9060094316@gmail.com

Varun Kumar P M
Bachelor of Computer Applications
Padmashree Institute of Management & Sciences, Bengaluru
varunrolexvarunrolex@gmail.com

Bharath C R
Bachelor of Computer Applications
Padmashree Institute of Management & Sciences, Bengaluru
bharathcr486@gmail.com

Jashwanth M K
Assistant Professor (Guide)
Padmashree Institute of Management & Sciences, Bengaluru
jashwanthgowda45@gmail.com

----------**********----------

Abstract:

The rapid growth of connected systems has widened the attack surface faced by organizations, while attackers increasingly rely on automation, obfuscation and zero-day techniques that signature-based defenses struggle to detect. Artificial Intelligence (AI), including Machine Learning (ML), Deep Learning (DL) and Reinforcement Learning (RL), offers a way to learn normal and malicious behavior directly from data and to respond at machine speed. This paper presents a structured review of AI techniques used in cybersecurity and of their applications in threat detection, covering intrusion detection, malware analysis, phishing and spam detection, fraud and insider-threat detection, and security-operations automation. A qualitative comparison between traditional and AI-based detection is provided across accuracy, adaptability, scalability, false-alarm behavior, explainability and resource cost. The paper also discusses benefits and limitations, including data quality, adversarial attacks, interpretability and privacy, and outlines future directions such as explainable AI, federated learning and adversarially robust models. The findings suggest that AI is most effective as a complement to, rather than a replacement for, traditional controls and human analysts.

Keywords — **Artificial intelligence, cybersecurity, machine learning, deep learning, intrusion detection, malware detection, anomaly detection, threat intelligence.**

----------**********----------

I. INTRODUCTION

Digital services now underpin finance, healthcare, government and everyday communication. As a result, cyberattacks such as ransomware, phishing, distributed denial of service and advanced persistent threats cause serious financial and operational damage. Security teams must monitor very large volumes of logs, network flows and endpoint events, and the number of alerts often exceeds what human analysts can investigate.

Traditional defenses rely mainly on signatures and hand-written rules. They are precise for known threats but weak against novel or modified attacks, and they require constant manual updates. AI-based methods address this by learning patterns from data, which allows them to generalize to

 ©IJSRED: All Rights are Reserved

unseen variants and to flag deviations from normal behavior [1], [3].

However, applying AI in security is not straightforward. Attackers can adapt, labeled data is scarce and often outdated, and operational environments are sensitive to false alarms [2]. This paper therefore examines both the potential and the limits of AI in cybersecurity.

Objectives

- Review the main AI techniques used in cybersecurity.
- Examine their application to threat detection tasks.
- Compare AI-based detection with traditional approaches.
- Identify benefits, limitations, challenges and future scope.

II. RESEARCH METHODOLOGY

This work follows a structured, qualitative literature-review methodology. No new experiments were run; the aim is to synthesize published evidence and to compare approaches on common criteria.

A. Scope and Sources

The review covers peer-reviewed journals, conference proceedings, standards documents and widely used textbooks on AI for security. Searches were carried out in IEEE Xplore, ACM Digital Library, SpringerLink, ScienceDirect and Google Scholar, using combinations of terms such as "machine learning", "deep learning", "intrusion detection", "malware detection", "anomaly detection" and "cybersecurity".

B. Selection Criteria

- Included: studies that apply AI/ML/DL to a defined security detection task and report methods, datasets or evaluation results.
- Excluded: purely opinion pieces, studies without a clear method, and duplicates.

Candidate papers were screened by title and abstract, then by full text, and the selected works were grouped by technique and by application area.

C. Analysis Framework

Each technique and application was compared on six criteria: detection capability on known and unknown threats, false-positive behavior, adaptability to change, scalability, explainability and resource requirements. Findings are summarized in Section VI.

III. LITERATURE REVIEW

Early surveys of data mining and ML for intrusion detection showed that classifiers such as decision trees, support vector machines and neural networks could detect attacks on benchmark data, but also warned that results on benchmarks may not carry over to real networks [1]. Sommer and Paxson argued that anomaly detection in operational settings is harder than in many other ML domains because of the high cost of errors, the diversity of normal traffic and the difficulty of evaluation [2]. Chandola et al. provided a general taxonomy of anomaly detection techniques that remains widely used in security research [3].

Deep learning surveys report that models such as convolutional, recurrent and autoencoder networks improve detection on several intrusion datasets, while noting the dependence on dataset quality and the age of commonly used benchmarks [4], [9]. More recent datasets such as CICIDS2017 attempt to reflect modern traffic and attack types [8]. Kitsune showed that an ensemble of autoencoders can perform online, lightweight anomaly detection on network streams [7]. Apruzzese et al. discussed the gap between laboratory results and deployment, including concept drift and adversarial manipulation [12].

Work on adversarial machine learning shows that small crafted perturbations can cause models to misclassify inputs [10], which is a direct concern for security systems that face active adversaries. In parallel, research on explainable AI, such as LIME, aims to make model decisions understandable to analysts [11]. Standards such as the NIST Cybersecurity Framework place detection within a wider risk-management process, which supports treating AI as one component of defense in depth [16].

Research Gap

Existing studies often evaluate a single technique on a single dataset. There is a need for a consolidated view that connects techniques to

 ©IJSRED: All Rights are Reserved

applications and compares them with traditional controls on operational criteria, which this paper provides.

IV. AI TECHNIQUES USED IN CYBERSECURITY

A. Supervised Learning

Supervised models learn from labeled examples of benign and malicious activity. Common choices include decision trees, Random Forests [5], Support Vector Machines [13], k-nearest neighbors and logistic regression. They perform well when representative labeled data is available and are widely used for malware classification, spam filtering and known-attack detection.

B. Unsupervised and Semi-Supervised Learning

Clustering, density-based methods, isolation-based methods and autoencoders model normal behavior and flag outliers [3], [7]. They do not require attack labels, which makes them suitable for detecting novel threats, but they tend to produce more false alarms when normal behavior varies widely.

C. Deep Learning

Convolutional Neural Networks (CNNs) are applied to byte-level malware images and traffic representations. Recurrent networks, particularly LSTMs [6], model sequences such as system-call traces, logs and network flows. Autoencoders learn compressed representations for anomaly scoring, and Transformer-based models are increasingly used for log and text analysis [4].

D. Reinforcement Learning

RL agents learn policies by interacting with an environment and receiving rewards [14]. In security, it is explored for adaptive response, automated penetration testing and dynamic defense, although most results remain at the research or simulation stage.

E. Natural Language Processing

NLP techniques analyze emails, URLs, threat reports and security logs. They support phishing detection, extraction of indicators of compromise from reports, and summarization of alerts for analysts.

F. Federated and Privacy-Preserving Learning

Federated learning trains a shared model across organizations or devices without centralizing raw data [17], which is attractive for security because logs often contain sensitive information.

TABLE I
TECHNIQUES AND TYPICAL SECURITY USES

Technique	Typical use	Main strength
Random Forest / SVM	Malware and spam classification	Strong with labeled data
Clustering / Isolation	Unknown-threat discovery	No labels needed
Autoencoder	Network anomaly scoring	Learns normal behavior
CNN	Malware images, traffic	Automatic features
LSTM / Transformer	Logs, sequences, text	Models context
Reinforcement learning	Adaptive response	Learns policies
Federated learning	Cross-site training	Data stays local

V. APPLICATIONS OF AI IN THREAT DETECTION

A. Network Intrusion Detection

AI models analyze flow features and packet statistics to detect scans, denial-of-service traffic, botnet communication and data exfiltration. Anomaly-based detectors complement signature-based systems by highlighting unusual flows [1], [7].

B. Malware Detection and Analysis

Static analysis uses features such as byte sequences, API imports and file structure, while dynamic analysis uses behavior observed in sandboxes. ML and DL models classify samples into families and can detect polymorphic or previously unseen variants better than exact-match signatures.

C. Phishing and Spam Detection

Classifiers use URL structure, page content, sender reputation and message text to detect phishing. NLP models help identify socially engineered messages that avoid known blacklists.

D. User and Entity Behavior Analytics

Behavior baselines are built for users, devices and services. Deviations such as unusual login

 ©IJSRED: All Rights are Reserved

times, abnormal data access or lateral movement can indicate compromised accounts and insider threats.

E. Fraud Detection

In payments and online services, models score transactions in near real time using amount, location, device and history, flagging suspicious activity for review.

F. Security Operations Automation

AI assists Security Operations Centers by correlating alerts, prioritizing incidents, enriching events with threat intelligence and suggesting or triggering response actions. This reduces alert fatigue and shortens response time, provided analysts remain in the loop.

G. Vulnerability and Threat Intelligence

NLP and ML help extract information from advisories and reports, estimate which vulnerabilities are more likely to be exploited, and cluster related campaigns.

VI. COMPARATIVE ANALYSIS: TRADITIONAL VS. AI-BASED DETECTION

Table II summarizes a qualitative comparison based on the surveyed literature. The ratings describe general tendencies; actual results depend on data, tuning and deployment.

TABLE II
TRADITIONAL VS. AI-BASED DETECTION

Criterion	Traditional (rules/signatures)	AI-based
Known threats	Very precise	Good; depends on training data
Unknown threats	Weak	Better at variants and anomalies
False positives	Low for exact rules	Can be higher without tuning
Adaptability	Manual updates	Retraining on new data
Scalability	Rule sets grow hard to manage	Handles large, high-dimensional data
Explainability	High; clear rule matched	Often low; needs XAI tools
Resources	Light	Needs data, compute, ML skills
Attack resistance	Evaded by small changes	Vulnerable to adversarial inputs

The comparison shows that the two approaches have complementary strengths. Rules provide precise, auditable detection of known threats, whereas AI extends coverage to variants and unusual behavior. In practice, layered architectures that combine both, with analyst review of AI findings, are the most reliable option [2], [12], [16].

VII. BENEFITS AND LIMITATIONS

A. Benefits

- Detection of previously unseen and modified attacks.
- Ability to process large volumes of data at high speed.
- Reduced manual effort through alert triage and automation.
- Continuous improvement as new data becomes available.
- Discovery of subtle patterns that are difficult to express as rules.

B. Limitations

- Dependence on large, representative and correctly labeled data.
- False positives and alert fatigue when models are poorly tuned.
- Vulnerability to adversarial examples and data poisoning [10].
- Limited interpretability of complex models [11].
- Concept drift: performance degrades as traffic and attacks change [12].
- Cost of compute, storage and specialist skills.
- Privacy and compliance concerns when handling sensitive logs.

VIII. CHALLENGES AND FUTURE SCOPE

A. Challenges

Benchmark datasets such as KDD-derived collections are old and contain known flaws [9], and even newer ones may not match a given environment [8]. Class imbalance is severe because attacks are rare compared with normal events. Attackers can probe and evade deployed models, and defenders must retrain without introducing instability. Regulatory and ethical requirements demand accountability for automated decisions, which is difficult with opaque models.

 ©IJSRED: All Rights are Reserved

Attackers also use AI themselves, for example to generate convincing phishing text or to automate reconnaissance, which raises the bar for defenders.

B. Future Scope

- Explainable AI that gives analysts understandable reasons for alerts.
- Adversarially robust training and continuous red-teaming of models.
- Federated and privacy-preserving learning across organizations.
- Online and continual learning to handle concept drift.
- Realistic, regularly refreshed public datasets and common evaluation protocols.
- Human-AI collaboration in security operations with measurable trust and oversight.
- AI-assisted defense for IoT, cloud and industrial control environments.

IX. CONCLUSION

AI has become an important part of modern cybersecurity. Supervised, unsupervised, deep and reinforcement learning methods extend detection beyond fixed signatures and support tasks ranging from intrusion and malware detection to phishing prevention and security-operations automation. Compared with traditional approaches, AI offers better coverage of unknown threats, greater scalability and adaptability, but at the cost of data requirements, higher false-alarm risk, limited explainability and exposure to adversarial attack.

The evidence reviewed suggests that the most effective strategy is a layered one in which AI augments rule-based controls and human expertise. Progress in explainability, robustness, privacy-preserving learning and realistic evaluation will determine how reliably AI can be trusted in operational defense.

REFERENCES

[1] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1153–1176, 2016.

[2] R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," in *Proc. IEEE Symp. Security and Privacy*, 2010, pp. 305–316.

[3] V. Chandola, A. Banerjee and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, no. 3, 2009.

[4] M. A. Ferrag, L. Maglaras, S. Moschoyiannis and H. Janicke, "Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study," *J. Inf. Secur. Appl.*, vol. 50, 2020.

[5] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[6] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[7] Y. Mirsky, T. Doitshman, Y. Elovici and A. Shabtai, "Kitsune: An ensemble of autoencoders for online network intrusion detection," in *Proc. Network and Distributed System Security Symp. (NDSS)*, 2018.

[8] I. Sharafaldin, A. H. Lashkari and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in *Proc. ICISSP*, 2018, pp. 108–116.

[9] M. Tavallaee, E. Bagheri, W. Lu and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in *Proc. IEEE Symp. Computational Intelligence for Security and Defense Applications*, 2009.

[10] I. J. Goodfellow, J. Shlens and C. Szegedy, "Explaining and harnessing adversarial examples," in *Proc. Int. Conf. Learning Representations (ICLR)*, 2015.

[11] M. T. Ribeiro, S. Singh and C. Guestrin, "Why should I trust you?": Explaining the predictions of any classifier," in *Proc. ACM SIGKDD*, 2016, pp. 1135–1144.

[12] G. Apruzzese, M. Colajanni, L. Ferretti, A. Guido and M. Marchetti, "On the effectiveness of machine and deep learning for cyber security," in *Proc. Int. Conf. Cyber Conflict (CyCon)*, 2018.

[13] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[14] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.

[15] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Pearson Education, 2021.

[16] National Institute of Standards and Technology, *Framework for Improving Critical Infrastructure Cybersecurity*, Version 1.1, 2018.

[17] B. McMahan, E. Moore, D. Ramage, S. Hampson and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. AISTATS*, 2017.

 ©IJSRED: All Rights are Reserved