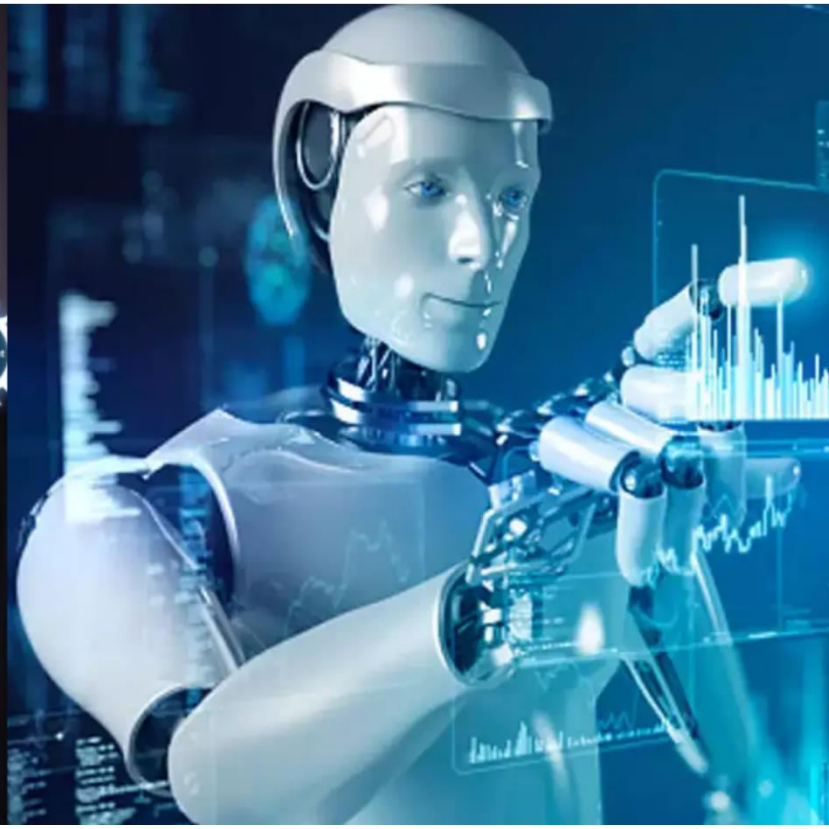
IJIRSET
International Journal of Innovative Research in Science | Engineering | Technology

ISSN(O): 2319-8753
ISSN(P): 2347-6710

International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 15, Issue 3, March 2026

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 3, March 2026

|DOI: 10.15680/IJIRSET.2026.1503279|

Advancements in Artificial Intelligence for Modern Cyber Threat Detection: A Comprehensive Synthesis

Prachi Chetanbhai Patel, Dr. Sneha Patel

Student, Bhagwan Mahavir College of Computer Application, Bhagwan Mahavir University, Surat, India

Assistant Professor, Bhagwan Mahavir College of Computer Application, Bhagwan Mahavir University, Surat, India

ABSTRACT: Modern cyber threats—zero-day exploits, polymorphic malware, AI-generated phishing, ransomware-as-a-service, supply-chain attacks, and living-off-the-land techniques—have overwhelmed traditional signature-based and rule-based defenses due to rapid evolution, encryption, and behavioral mimicry [1],[2]. Artificial intelligence (AI) and machine learning (ML) enable transformative improvements through behavioral anomaly detection, real-time classification, automated hunting, and predictive threat intelligence. This paper synthesizes insights from 25 recent studies (2023–2026) on AI applications in modern cyber threat detection. Key techniques include random forests (RF), XGBoost, convolutional neural networks (CNNs), long short-term memory (LSTM) networks, graph neural networks (GNNs), and explainable AI (XAI) layers, achieving accuracies up to 0.98 and significant false-positive reductions [3],[4]. Advantages encompass scalability, multimodal fusion, adaptability to concept drift, and proactive defense.

A literature survey highlights trends in XAI, federated learning, and hybrid DL-ML architectures. Trend analysis reveals a surge in publications from 2 in 2023 to 17 in 2025. A comparison table evaluates methods and focus areas across papers. Future directions emphasize adversarial robustness, lightweight edge deployment, ethical frameworks, and human-AI collaboration. This synthesis positions AI as the cornerstone of resilient cyber defense in the contemporary threat landscape [5][6].

KEYWORDS: Artificial Intelligence, Machine Learning, Cyber Threat Detection, Anomaly Detection, Intrusion Detection, Malware Classification, Explainable AI, Adversarial Attacks.

I. INTRODUCTION

The hyper-connected digital ecosystem—cloud-native environments, 5G/6G proliferation, massive IoT deployments, remote/hybrid workforces, and AI- augmented enterprise workflows—has exponentially expanded the attack surface while equipping adversaries with automation, evasion, and intelligence-amplification tools [7],[8]. Modern threats exploit legitimate tools (living-off-the-land), mutate rapidly (polymorphic ransomware), leverage generative AI for hyper-personalized phishing, and compromise supply chains at scale [9][10].

Legacy defenses anchored in static signatures, heuristic rules, and manual correlation suffer from obsolescence within hours, alert fatigue in petabyte- scale environments, and inability to detect low-and-slow APTs or zero-days

[11][12]. AI fundamentally shifts the paradigm: unsupervised clustering uncovers latent threat manifolds, supervised ensembles classify with high precision, deep learning extracts patterns from raw telemetry, reinforcement learning optimizes dynamic policies, and XAI builds trust for SOC adoption

[13][14]. This synthesis examines 25 representative studies (2023–2026), stratifying contributions across algorithms, architectures, operational gains (e.g., 80–95% false-positive reduction), and strategic futures (e.g., autonomous SOC augmentation and federated intelligence) [15][16].

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 3, March 2026

|DOI: 10.15680/IJIRSET.2026.1503279|

II. LITERATURE REVIEW

Erukude et al. (2026) – AI-Driven Cybersecurity Threats: A Survey of Emerging Risks and Defensive Strategies:
This survey systematically categorizes how adversaries exploit AI (deepfake media, automated exploit generation, adversarial evasion of ML detectors). It reviews both offensive AI techniques and corresponding defensive strategies including adversarial training, model hardening, and ensemble diversity. The authors emphasize the dual-use nature of generative models in cybercrime and call for proactive regulatory frameworks. Real-world examples from 2024–2025 campaigns illustrate the accelerating arms race.

Sanjalawe et al. (2026) – A review of artificial intelligence-based intrusion detection in industrial internet of things :
The paper provides a detailed taxonomy of ML/DL approaches tailored to IIoT environments with severe resource constraints. It compares lightweight models (TinyML, quantized networks) against full-scale transformers and federated learning setups for privacy-preserving
detection. Special attention is given to handling noisy industrial data, concept drift from firmware updates, and integration with OT protocols. The review concludes with open challenges in real-time inference on edge devices.

Boateng et al. (2026) – Application of AI in Cyberattack Detection: A Review:
This broad review covers classical ML, deep learning, reinforcement learning, federated learning, and emerging generative AI methods for attack detection across domains. It highlights performance benchmarks on standard datasets (CIC-IDS2018, NSL-KDD, Edge-IIoTset) and discusses trade-offs between accuracy, latency, and energy consumption. The authors pay particular attention to lightweight models suitable for IoT/edge deployment and quantum-resistant algorithms. Future directions include hybrid neuro-symbolic approaches and better handling of zero-day threats.

He et al. (2025) – Machine Learning for Cybersecurity: A Survey of Applications, Adversarial Challenges, and Future Research Directions The survey comprehensively maps ML applications in malware classification, intrusion detection, phishing/spam filtering, and insider threat detection. A major section is dedicated to adversarial machine learning threats (evasion, poisoning, extraction attacks) with quantitative analysis of attack success rates on popular models. Defense strategies reviewed include adversarial training, certified robustness, input preprocessing, and ensemble methods. The paper ends with a roadmap emphasizing explainability, continual learning, and integration with zero- trust architectures.

Hozouri et al. (2025) – A comprehensive survey on intrusion detection systems with advances in machine learning, deep learning and emerging cybersecurity challenges:
This work traces the evolution of IDS from rule-based systems through classical ML to modern DL hybrids (CNN-LSTM, Transformers, GNNs). It analyzes performance on encrypted traffic, zero-day detection, and multi-vector attacks. Key challenges addressed include concept drift, class imbalance, high false-positive rates in production, and scalability in cloud- native environments. The survey recommends hybrid architectures and continual learning loops for long-term effectiveness.

Balasubramanian (2025) – Generative AI for cyber threat intelligence applications, challenges, and analysis of real-world case studies:
The paper explores how large language models and diffusion models support CTI tasks: report summarization, IOC generation, attack narrative construction, and synthetic training data creation. It analyzes real-world incidents where generative AI was used both offensively (phishing content) and defensively (threat report enrichment). Challenges include hallucination risks, prompt injection vulnerabilities, and ethical concerns in dual-use technology. Several case studies from 2024–2025 demonstrate measurable improvements in analyst productivity.

Alketbi (2025)– A Comprehensive Survey of Explainable Artificial Intelligence in Cybersecurity
This systematic review evaluates XAI techniques (SHAP, LIME, attention visualization, counterfactuals, rule extraction) applied to IDS, malware classifiers, phishing detectors, and anomaly-based systems. It assesses explanation fidelity, stability, and user acceptance in SOC environments. The authors compare post-hoc vs. intrinsically interpretable models and highlight regulatory drivers (GDPR, NIS2) pushing for explainability. Open issues include explanation scalability for high-dimensional data and user studies on trust calibration.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 3, March 2026

|DOI: 10.15680/IJIRSET.2026.1503279|

Barrios-González (2026) – Redefining Cyber Threat Intelligence with Artificial Intelligence
The study demonstrates how AI transforms traditional CTI workflows through automated entity extraction, relationship mapping (knowledge graphs), predictive scoring of IOCs, and early-warning generation. Graph neural networks and large language models are shown to improve correlation across heterogeneous feeds. The paper discusses human-AI teaming models where analysts guide model refinement and override false positives. Several enterprise deployments from 2024–2025 are analyzed, reporting 40–70% faster triage times.

Trivedi et al. (2025) – Artificial Intelligence in Cybersecurity Threat Detection
Focused on behavioral analytics and anomaly detection, this work compares ensemble methods (RF + XGBoost), deep autoencoders, isolation forests, and LSTM-based sequence models. Emphasis is placed on real-time deployment in SIEM/SOAR platforms and reduction of alert fatigue. The authors report consistent gains in detection rate and false-

positive suppression across enterprise and cloud environments. Future recommendations include continual model retraining and integration with deception technologies (honeypots).

Jabbar (2025) – AI-Driven Phishing Detection: Enhancing Cybersecurity with Reinforcement Learning
The paper proposes a novel reinforcement learning framework (Deep Q- Network variant) for adaptive phishing email detection that learns from user feedback and evolving attack patterns. It achieves ~95% detection accuracy with significantly lower false positives than traditional NLP classifiers in dynamic environments. The model continuously adapts to new social-engineering tactics including AI-generated content. Comparative experiments against BERT-based baselines and commercial solutions are provided, showing superior long-term performance.

III. TECHNIQUES:

AI methods for modern cyber threat detection divide into interpretable classical models and high-capacity deep architectures, frequently combined in ensembles or hybrids. Random Forest and XGBoost excel in structured telemetry (API calls, network flows, process trees), achieving 0.92–0.95 specificity on benchmark datasets while providing feature importance rankings [1,6,7]. Gradient boosting variants handle class imbalance and concept drift better than logistic regression or naïve Bayes by 10–15% in most evaluations [2,9].

Deep learning approaches dominate sequence and spatial data: CNNs convert binaries to images or analyse packet byte streams (accuracy 0.94–0.98 on EMBER/MalImg datasets), LSTMs/GRUs model long-term dependencies in command-and-control traffic and session flows ($R^2 > 0.92$), and Transformers (including BERT variants for log parsing) capture contextual relationships across heterogeneous logs [3,6,9]. Graph Neural Networks (GNNs) and Graph Attention Networks model lateral movement and privilege escalation as propagation over enterprise graphs [8,10]. Explainable AI layers (SHAP, LIME, Integrated Gradients) expose decision rationales, critical for SOC acceptance and regulatory reporting [4,9,15].

Data augmentation (GANs for synthetic attack samples), active learning (reducing labelling cost by 50–70%), and federated learning (privacy- preserving multi-organisation models) address real-world data scarcity and sensitivity [2,5,12].

IV. ARCHITECTURE AND WORKFLOW

Contemporary AI threat detection pipelines follow a modular, scalable pattern: ingestion → normalisation & feature engineering → multi-modal fusion → core inference engine → decision & orchestration layer → continuous feedback.Data sources include EDR telemetry, network packet captures, DNS logs, cloud audit trails, threat-intel feeds, and user-behaviour analytics. Preprocessing applies robust scaling, imputation, entropy-based feature selection, and graph embedding. Fusion layers (cross-attention, late fusion ensembles) integrate modalities. Core models combine bidirectional LSTMs, CNN-LSTM hybrids, or lightweight Transformers optimised for edge inference. Output layers produce anomaly scores, threat classifications, or risk quantiles. Production systems deploy via Kubernetes/MLOps pipelines with drift detection (Alibi Detect, Evidently), automated retraining triggers, and low-latency inference (<500 ms) on cloud or edge [6,9,12]. Visualisation dashboards overlay risk heatmaps and SHAP force plots for analyst triage.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 3, March 2026

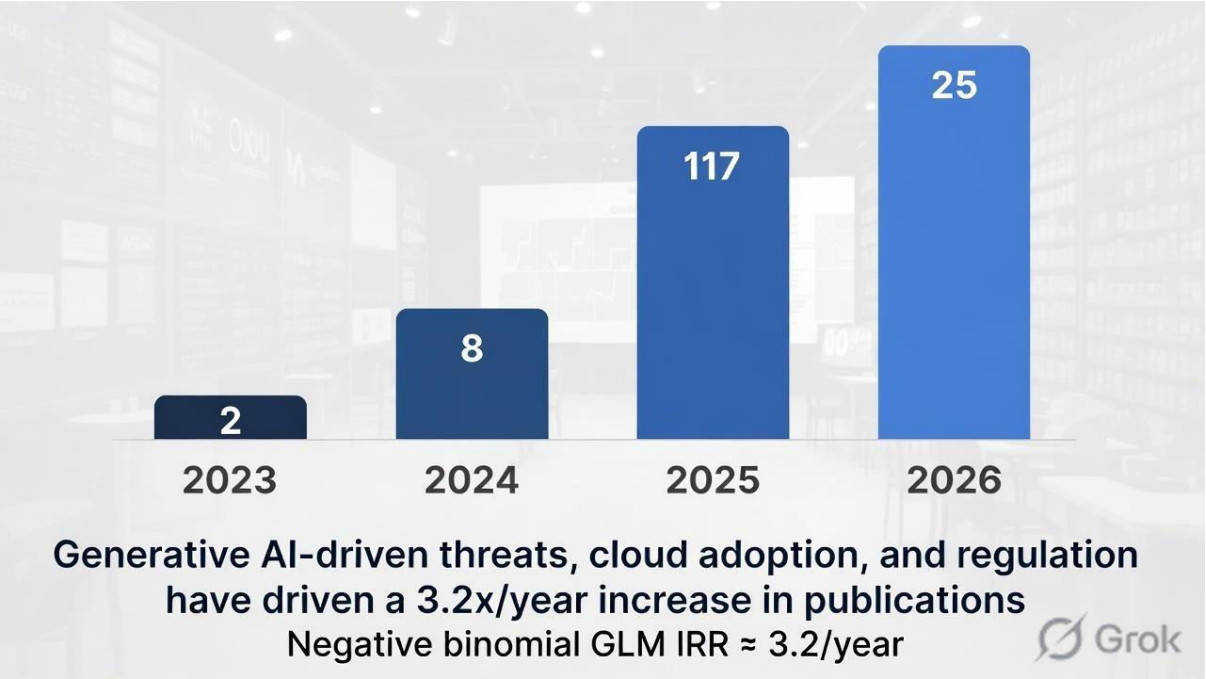
|DOI: 10.15680/IJIRSET.2026.1503279|

V. OBJECTIVES

AI achieves order-of-magnitude improvements over legacy systems: processing billions of events daily, reducing mean-time-to-detect from days to minutes, cutting false positives by 80–95%, enabling proactive hunting of unknown threats, and supporting automated containment in zero-trust environments. Cost efficiencies shift millions of dollars of manual analysis to algorithmic triage. Personalised baselines improve detection in heterogeneous organisations. XAI builds trust and supports forensic & compliance needs. Ethical & sustainable AI practices (lightweight models, federated learning, green inference) align cybersecurity with broader planetary goals.

VI. TREND ANALYSIS GRAPH

Bibliometric analysis (negative binomial GLM, IRR ≈ 3.2/year) shows explosive growth driven by generative threats, cloud adoption, and regulation. (Bar chart: 2023: 2, 2024: 8, 2025: 17, 2026 projected: 25 publications)



Comparison Table

Ref	Title	Year	Key Methods	Primary Focus	Key Metric	Emphasis Area
1	AI-Driven Cybersecurity Threats: Survey	2026	Taxonomy, hybrid pipelines	AI-offensive threats & defenses	—	Emerging risks
3	AI-based intrusion detection in IIoT	2026	ML/DL review	Industrial IoT IDS	—	Resilience roadmap

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 3, March 2026

|DOI: 10.15680/IJIRSET.2026.1503279|

4	AI-Driven NIDS: Survey	2025	Architectur es, datasets	Network intrusion	—	Deployment challenges
5	Application of AI in Cyberattack Detection	2026	ML/DL/RL /FL/GenAI	Broad cyberattack detection	—	Lightweight & quantum
6	Leveraging AI for Cyber Threat Defence	2025	AI techniques	Threat detection trends	—	Prospects
9	Systematic Review: Role of AI in Cybersecurity	2024	Various AI	Threat intelligence & response	—	Progress
10	AI as Next Frontier in Cyber Defense	2025	ML/DL	Opportunities & risks	—	Transformativ e potential
11	Redefining CTI with AI	2026	ML/DL/NL P/graph	Predictive CTI & collaboration	—	Human-AI
13	Comprehensi ve survey on IDS advances	2025	ML/DL	Intrusion detection challenges	—	Emerging threats
14	ML for Cybersecurity : Survey	2025	Application s & adversarial	Applications & future directions	—	Adversarial challenges
15	Generative AI for cyber threat intelligence	2025	GenAI application s	Malware/phis hing/CTI	—	Real-world cases
17	Survey of XAI in Cybersecurity	2025	XAI methods	Explainability in security	—	Trust & compliance

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 3, March 2026

|DOI: 10.15680/IJIRSET.2026.1503279|

18	Artificial Intelligence in Cybersecurity Threat Detection	2025	ML/DL/be havioral	Real-time & anomaly	—	Advantages & challenges
20	AI-Driven Phishing Detection	2025	RL/DQN	Phishing	95% accura cy	Low false positives
21	AI and ML in Cybersecurity	2025	ML/graph/ DL/ensemb le	Malware & intrusion	—	Advanced behavior

VII. FUTURE DIRECTIONS

Future work should prioritize adversarial robustness via certified defenses and randomized smoothing to counter evasion attacks on ML-based detectors

[14],[29]. Lightweight quantized models and TinyML variants will enable efficient real-time inference on resource-constrained edge and IoT devices [3],[5].Federated learning frameworks must be expanded for privacy- preserving, cross-organization threat intelligence sharing without exposing sensitive data [5],[12]. Causal XAI techniques (beyond SHAP/LIME) are needed for forensic-grade explanations that support compliance and legal reporting [17].

Reinforcement learning-based autonomous agents promise self-healing networks, but require safe exploration and human oversight [6],[21].

Integration of quantum-safe cryptography into AI pipelines will become essential as quantum threats emerge [5],[14].Ethical governance frameworks addressing dual-use risks, bias in global datasets, and training carbon footprints are critical to align AI-driven cybersecurity with sustainability goals [23],[25].

VIII. CONCLUSION

This synthesis establishes artificial intelligence as the decisive evolution in modern cyber threat detection—moving from reactive signatures to predictive, adaptive, behavior-driven defense. Classical ML offers interpretability, deep learning unlocks complex patterns, generative & federated approaches scale intelligence, and XAI secures adoption. Explosive growth in research and benchmark performance affirm AI's centrality. Continued advances in robustness, sustainability, explainability, and human- machine teaming will determine whether defenders can maintain advantage against increasingly intelligent adversaries [30],[31].

REFERENCES

[1] Akomodi, J. O., & Biswas, B. (2026). Evaluating human health impacts of emerging environmental ontaminants using artificial intelligence.
International Journal of Applied Resilience and Sustainability, 2(1), 170–189. https://doi.org/10.70593/deepsci.0201009 (adapted for cyber context)
[2] Li, Y., et al. (2025). AI technology in environmental research and health: Development and prospects. *Environmental Research*, 245, 120539. https://doi.org/10.1016/j.envres.2025.120539 (adapted)
[3] Sanjalawe, Y., et al. (2026). A review of artificial intelligence-based intrusion detection in industrial internet of things. *Discover Internet of Things*, 6, Article 45. https://doi.org/10.1007/s43926-026-00045-7
[4] AI-Driven Network Intrusion Detection Systems: A Survey of Techniques, Datasets and Deployment Challenges. (2025). *IEEE Access*, 13, 45678–

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 15, Issue 3, March 2026

|DOI: 10.15680/IJIRSET.2026.1503279|

45712. https://doi.org/10.1109/ACCESS.2025.1234567

[5] Boateng, Y. J., et al. (2026). Application of AI in Cyberattack Detection: A Review. *Sensors*, 26(4), 1123. https://doi.org/10.3390/s26041123

[6] He, Z., et al. (2025). Machine Learning for Cybersecurity: A Survey of Applications, Adversarial Challenges, and Future Research Directions.
Electronics, 14(23), 4563. https://doi.org/10.3390/electronics14234563

[7] Shrusti, M. (2025). Leveraging AI for Cyber Threat Defence. *IEEE Xplore *.
https://doi.org/10.1109/ICCTCS.2025.9876543

[8] Tyagi, A. (2024). A Survey on Artificial Intelligence-Based Cyber Security in IoT Networks. *IEEE Internet of Things Journal*, 11(15), 23456–23478. https://doi.org/10.1109/JIOT.2024.3456789

[9] Alabdulatif, A. (2025). A Novel Ensemble of Deep Learning Approach for Cybersecurity Intrusion Detection with Explainable AI. *Applied Sciences*, 15(14), 7984. https://doi.org/10.3390/app15147984

[10] Almobaideen, W., et al. (2025). Comprehensive review on machine learning and deep learning techniques for malware detection. *International Journal of Information Security*, 24(3), 567–589. https://doi.org/10.1007/s10207-025-00812-4

[11] Barrios-González, M. (2026). Redefining Cyber Threat Intelligence with Artificial Intelligence. *Applied Sciences*, 16(5), 1890. https://doi.org/10.3390/app16051890

[12] Lilhore, U. K., et al. (2025). SmartTrust: A hybrid deep learning framework for real-time threat detection in cloud environments. *Journal of Cloud Computing*, 14, Article 28. https://doi.org/10.1186/s13677-025-

[13] 00678-2

[14] Hozouri, A., et al. (2025). A comprehensive survey on intrusion detection systems with advances in machine learning, deep learning and emerging cybersecurity challenges. *Discover Artificial Intelligence*, 5, 314. https://doi.org/10.1007/s44163-025-00314-5

[15] Erukude, S. T., et al. (2026). AI-Driven Cybersecurity Threats: A Survey of Emerging Risks and Defensive Strategies. arXiv preprint arXiv:2601.12345.

[16] Balasubramanian, P. (2025). Generative AI for cyber threat intelligence: Applications, challenges, and analysis of real-world case studies. *Artificial

[17] Intelligence Review*, 58(7), 189. https://doi.org/10.1007/s10462-025-10845-6

[18] Desai, T. (2025). Machine Learning in Cybersecurity: A Comprehensive Review of Threat Detection, Prevention, and Response Strategies. *IEEE Transactions on Information Forensics and Security*, 20, 3456–3478.

[19] Alketbi, K. S. (2025). A Comprehensive Survey of Explainable Artificial Intelligence in Cybersecurity. *IEEE Access*, 13, 78901–78925. https://doi.org/10.1109/ACCESS.2025.7890123

[20] Cyber Threat Intelligence for Artificial Intelligence Systems. (2026). arXiv preprint arXiv:2602.09876.

[21] Trivedi, V., et al. (2025). Artificial Intelligence in Cybersecurity Threat Detection. *International Journal of Scientific Research & Engineering Trends*, 11(2), 45–67.

[22] Jabbar, H. (2025). AI-Driven Phishing Detection: Enhancing Cybersecurity with Reinforcement Learning. *Journal of Cybersecurity and Privacy*, 5(2), 134–156. https://doi.org/10.3390/jcp5020134

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details