Literature Review: Artificial Intelligence Applications in Cybersecurity - Threat Detection, Anomaly Detection, and System Integrity

1. Introduction

In the rapidly evolving digital landscape, the frequency and sophistication of cyberattacks have outdone traditional security frameworks. Recent literature strongly supports the integration of Artificial Intelligence (AI) into cybersecurity to meet the demands of this complex threat environment. AI technologies in particularl, Machine Learning (ML), Deep Learning (DL), Reinforcement Learning (RL), and Natural Language Processing (NLP), are increasingly used to enhance threat detection, automate response, and fortify system resilience. This literature review synthesizes insights from recent peer-reviewed works including those by Balaji et al. (2024), Singh (2025), IJIRSS (2025), and Lawrence (2025), offering a consolidated analysis of the state of AI in cybersecurity, its technical contributions, and associated challenges.

2. Core Applications of AI in Cybersecurity

2.1 Threat Detection

AI's foremost application is in real-time threat detection, where supervised learning models analyze vast data streams including network packets, user behavior logs, and malware signatures. As highlighted by Balaji et al. (2024) and IJIRSS (2025), ML models like Random Forests and SVMs have improved threat detection accuracy by 17–35% compared to legacy systems. Singh (2025) emphasizes agentic AI's autonomous threat recognition, particularly for rapid incident response, while Lawrence (2025) further explores predictive threat intelligence using AI to preemptively identify emerging threats.

2.2 Anomaly Detection

Unsupervised and semi-supervised models, especially autoencoders and deep neural networks, are central to anomaly detection strategies. Lawrence (2025) illustrates the use of DL in identifying subtle behavioral deviations often missed by rule-based systems - crucial for catching insider threats and zero-day exploits. Balaji et al. (2024) describe this function within AI-driven SOCs that continuously refine detection logic using incoming telemetry.

2.3 Autonomous Incident Response

Reinforcement Learning (RL) and agentic AI offer self-directed response systems capable of mitigating threats autonomously. Singh (2025) documents RL models that execute countermeasures within seconds of detecting anomalies. IJIRSS (2025) identifies RL's role in automated defense orchestration, especially in large-scale SOCs. However, the ethical implications of these autonomous systems—particularly accountability and control—remain underexamined (Lawrence, 2025).

2.4 Securing AI Systems

AI models themselves are vulnerable. Lawrence (2025) identifies risks such as adversarial attacks, data poisoning, and model inversion. To mitigate these, cutting-edge methods like homomorphic

encryption and federated learning are proposed for privacy-preserving model training and threat analysis.

3. Cross-Technology Evaluation and Metrics

Evaluation across studies centers on metrics such as accuracy, precision, recall, false positive rate, system overhead, and CPU utilization. Notably, DL and RL consistently lead in detection accuracy and threat response speed. Federated learning and privacy-preserving computation reduce exposure risks during model training. However, false positives and lack of interpretability remain persistent across AI models, limiting their utility in high-stakes decision-making.

4. Emerging Research Themes

4.1 Explainable and Trustworthy AI

There is increasing demand for Explainable AI (XAI) in cybersecurity to ensure transparency and trust. Singh (2025) and IJIRSS (2025) emphasize the need for interpretable models to assist human analysts. Lawrence (2025) proposes integrating regulatory compliance mechanisms to govern AI deployment.

4.2 Hybrid and Scalable Frameworks

Combining multiple AI models like ML + RL, DL + NLP, is emerging as a strategy for enhancing resilience and minimizing false positives. IJIRSS (2025) and Lawrence (2025) both advocate for hybrid intelligent systems capable of scaling horizontally and vertically. Singh (2025) highlights adaptive learning loops that allow agentic AI systems to evolve with emerging threats.

4.3 Ethics, Governance, and Policy

AI's autonomous nature in cybersecurity introduces regulatory, ethical, and operational dilemmas. Concerns over autonomy without accountability, data privacy, and human oversight are flagged across studies. Transparent frameworks and global standards are needed to ensure AI is safe, fair, and auditable.

5. Summary of Key Insights

A comparative summary across application domains shows threat detection accuracy improved with ML/DL models, anomaly detection is crucial for zero-day and insider threats, autonomous response ensures faster mitigation through RL and agentic AI, and the importance of privacy-preserving and resilient architectures in securing AI.

Table 1. key insights summary

Domain	Contribution of AI	Key Techniques	Challenges
Threat Detection	Improved accuracy & speed	ML (SVM, RF), DL (CNN, LSTM)	False positives, model opacity
Anomaly Detection	Detect zero-day, insider threats	Autoencoders, LSTM, Unsupervised DL	High complexity, resource consumption
Automated Response	Real-time mitigation	RL, Agentic AI	Trust, accountability
Securing AI	Protection of AI assets	Homomorphic encryption, Federated Learning	Model attacks, regulatory gaps

6. Conclusion and Future Directions

Across multiple domains, AI has shifted cybersecurity from reactive to proactive and autonomous operations. As AI tools become more embedded in critical infrastructure, focus must shift toward explainable models, adversarial robustness, and responsible regulation. Ongoing collaboration between researchers, practitioners, and policymakers will be essential to realizing the full potential of AI in cybersecurity.

References

1. Balaji, T.S., et al. (2024). Research on the Application of Artificial Intelligence in Cybersecurity. Integrating Advanced Technologies to Improve Threat Detection and Response. IEEE.

2. Singh, A. (2025). Agentic AI for Cybersecurity: Threat Detection and Response. ResearchGate.

3. IJIRSS. (2025). AI Integration in Cybersecurity Software: Threat Detection and Response. International Journal of Innovative Research and Scientific Studies, 8(3), 3907–3921.

4. Lawrence, E. (2025). AI in Cybersecurity: Threat Detection, Anomaly Detection, and Secure AI Systems. ResearchGate.