Volume 10, Issue 5, May – 2025 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25may2116

Advancing Intelligent Threat Detection Systems Powered by AI: A Comprehensive Review and Conceptual Framework

[1]Almash Saifi; [2]Mukul Sharma; [3]Mragesh Pratap Singh

Department of Computer Application, Swami Vivekanand Subharti University, Meerut

Publication Date: 2025/06/09

Abstract: The development of intelligent threat detection systems is required due to the increasing complexity and frequency of cyber attacks. With the use of sophisticated anomaly detection and behavioural analysis, artificial intelligence (AI) has become a crucial element in improving network security. With an emphasis on AI techniques applicable to network behavior analysis, current machine learning algorithms for anomaly detection, theoretical risk evaluation of institutional network threats, and best practices for deploying AI-driven detection systems in real-world networks, this paper provides an extensive review of recent literature on AI-driven threat detection. Additionally, by incorporating knowledge from recent studies and business procedures, we offer a conceptual architecture for an AI-based threat detection system.

How to Cite: Almash Saifi; Mukul Sharma; Mragesh Pratap Singh (2025) Advancing Intelligent Threat Detection Systems Powered by AI: A Comprehensive Review and Conceptual Framework. *International Journal of Innovative Science and Research Technology*, 10(5), 4096-4099. https://doi.org/10.38124/ijisrt/25may2116

I. INTRODUCTION

The digital transformation across industries has led to an expanded attack surface, making networks more susceptible to sophisticated cyber threats. Traditional security measures often fall short in detecting and mitigating these evolving threats. AI offers promising solutions by enabling proactive threat detection through pattern recognition, anomaly detection, and predictive analytics. This paper aims to explore the current state of AI in threat detection and propose a structured framework for its implementation.

II. AI TECHNIQUES IN NETWORK BEHAVIOR ANALYSIS

AI techniques have revolutionized network behavior analysis by enabling real-time monitoring and anomaly detection. Machine learning algorithms, such as Support Vector Machines (SVM), Decision Trees (DT), and Artificial Neural Networks (ANN), have been employed to classify network traffic and identify malicious activities. For instance, a study demonstrated that DT achieved an F1-score of 99.96% and an AUC of 99.93% in intrusion detection tasks, outperforming SVM and ANN models.

Deep learning approaches, including Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, have also been utilized for their ability to capture complex patterns in network data. Moreover, hybrid models combining supervised and unsupervised learning techniques have shown improved detection accuracy and reduced false positives.

III. MACHINE LEARNING ALGORITHMS FOR ANOMALY DETECTION

Anomaly detection is crucial for identifying deviations from normal network behavior that may indicate security breaches. Various machine learning algorithms have been explored for this purpose:

- Random Forest (RF) and XGBoost: These ensemble methods have demonstrated high detection accuracy in malware identification tasks.
- Isolation Forest: Effective in detecting anomalies by isolating outliers in the data.
- Deep Reinforcement Learning (DRL): Utilized for adaptive threat detection, DRL models can learn optimal defense strategies in dynamic environments.

The integration of these algorithms into intrusion detection systems enhances their capability to detect and respond to emerging threats.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25may2116

IV. CONCEPTUAL ARCHITECTURE FOR AI-BASED THREAT DETECTION FRAMEWORK

Based on the literature review, we propose a conceptual architecture for an AI-based threat detection framework comprising the following components:

- Data Collection Layer: Aggregates data from various sources, including network traffic, system logs, and user activities.
- Preprocessing Layer: Cleanses and normalizes data to ensure quality and consistency.
- Feature Extraction Layer: Identifies relevant features using techniques like Principal Component Analysis (PCA) and autoencoders.
- Detection Engine: Employs machine learning and deep learning models to detect anomalies and classify threats.
- Decision Support System: Provides actionable insights and recommendations for threat mitigation.
- Feedback Loop: Incorporates feedback to continuously improve model accuracy and adapt to new threats.

This architecture emphasizes modularity, scalability, and adaptability to cater to diverse organizational needs.

V. ASSESSMENT OF THEORETICAL RISK AND SIMULATION OF INSTITUTIONAL NETWORK DANGERS

Evaluating risks entails determining the possibility and possible consequences of challenges to institutional networks. AI makes this possible by simulating assault scenarios and using predictive modeling. To improve machine learning models' performance in anomaly detection tasks, for instance, Bayesian optimization approaches have been used.

The application of adversarial machine learning also emphasizes the necessity of strong models that are resistant to efforts to trick AI systems. A thorough grasp of possible vulnerabilities is ensured by including such factors into risk models.

VI. GUIDELINES FOR IMPLEMENTING AI-DRIVEN DETECTION SYSTEMS IN REAL-WORLD NETWORKS

Implementing AI-driven detection systems requires careful planning and consideration of various factors:

A. Data Quality

➢ Significance: Any effective AI-driven detection system is built on top of high-quality data. Low trust, high false positive/negative rates, and faulty models are all consequences of poor data.

- The level of representation: The dataset has to capture the variety of protocols, traffic kinds, and attack vectors that occur in the actual world.
- Labeling: Properly labeled data is necessary for supervised models. Mislabeling, such as when harmless traffic is mistakenly classified as harmful, can significantly lower model accuracy.
- Volume and Velocity: A significant amount of data is needed for AI models, particularly deep learning models. Real-time data processing and ingestion must be managed by systems.
- Noise Reduction: Filtering redundant, unnecessary, or noisy data (such as background traffic that isn't helpful for threat detection) should be part of preprocessing.
- Data Augmentation: When actual data is limited, use simulations, anomaly injection, or synthetic data to increase and diversity training sets.

B. Model Selection

➢ *Significance: Selecting the appropriate AI/ML model guarantees conformity with the type of network traffic, computing limitations, and detection objectives.*

- Anomaly Detection vs. Signature-Based Detection: o Models based on anomalies, such as Auto-encoder and Isolation Forests, are effective against zero-day attacks.

Known attack patterns are effectively handled by signature-based models (such as rule-based classifiers).

➢ *Types of Models:*

- Traditional machine learning (e.g., Random Forest, SVM): Effective for data from tabular networks.
- Deep Learning (e.g., CNN, LSTM): Fits well with sequence and temporal data (e.g., log time-series).
- Graph-based machine learning: To identify linkages or lateral movement across systems.
- Environment Matching: Heavy models for cloud or data center installations vs lightweight models for edge detection.

C. Scalability

➢ *Significance: Without sacrificing speed, the detection system must handle expanding data quantities and network complexity.*

- Modular Architecture: To enable flexible scalability, use microservices and containerization (such as Docker and Kubernetes) in your design.
- Distributed Processing: For real-time data stream processing, use frameworks such as Apache Kafka, Spark, or Flink.
- Edge vs. Centralized Processing: Strike a balance between central aggregation (for comprehensive threat correlation) and local processing (for speed and privacy).

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25may2116

- Load balancing: To prevent bottlenecks, use intelligent data routing and load distribution models.

D. Explainability

➢ Significance: In order to satisfy compliance standards, win over stakeholders, and enable human analyst involvement, explainability is essential.

- Explainable AI (XAI) Methods:
- SHAP and LIME: Describe the predictions of each model separately. Determine which features influence decisions by doing a feature importance analysis.
- Visual Interfaces: Dashboards that provide confidence levels, anomaly ratings, and decision justification.
- Model Transparency: Whenever feasible, use models that are naturally interpretable, such as decision trees. Enable human analysts to review and improve model decisions through feedback loops.

E. Continuous Learning

➢ *Significance: Threat landscapes evolve rapidly; static models become obsolete. Continuous learning ensures adaptability and long-term effectiveness.*

Considerations:
- Online Learning Algorithms: Adaptive neural networks and Hoeffding trees are examples of online learning algorithms that gradually update models in response to incoming input.
- Drift Recognition: Track alterations in statistical characteristics over time, such as data or idea drift, and initiate retraining.
- Retraining Pipelines: Use CI/CD for ML (MLOps) to automate data labeling, model assessment, and deployment cycles.
- Feedback Integration: To enhance learning, include input from human analysts into the training cycle.

F. Compliance and Ethics

➢ *Significance: AI systems that handle network data must respect privacy, follow the law, and refrain from acting in an immoral or discriminating manner.*

- Data Privacy Laws: Make sure that data collection, retention, and processing adhere to laws such as the CCPA, GDPR, HIPAA, and others.
- Anonymization: PII (personally identifiable information) in datasets is eliminated or hidden.
- Audit Trails: Keep track of AI decision logs for reporting compliance and responsibility.
- Verify that models do not unjustly discriminate against certain users, devices, or traffic types in order to mitigate bias.
- Frameworks for Responsible AI: Adhere to best practices from frameworks such as IEEE Ethically Aligned Design, ISO/IEC 23053, and NIST AI RMF.

VII. CONCLUSION

AI has significantly advanced the capabilities of threat detection systems, offering proactive and adaptive security solutions. By leveraging machine learning and deep learning techniques, organizations can enhance their ability to detect and respond to cyber threats. The proposed conceptual framework and implementation guidelines serve as a foundation for developing robust AI-based security systems. Future research should focus on improving model explainability, resilience against adversarial attacks, and integration with broader cybersecurity strategies. Real-world network implementation of AI-driven detection systems presents a complex technical and governance problem. In addition to technological expertise, operational maturity, regulatory congruence, and ethical foresight are critical components of success. A well-designed system will be strong, transparent, and flexible enough to change with new threats and expanding infrastructures.

REFERENCES

[1]. Alqahtani, A., & AlShaher, H. (2024). Anomaly-Based Intrusion Detection Systems Using Machine Learning. Journal of Cybersecurity and Information Management, 14(1), 20-33.
[2]. Kimanzi, R., Kimanga, P., Cherori, D., & Gikunda, P. K. (2024). Deep Learning Algorithms Used in Intrusion Detection Systems -- A Review. arXiv preprint arXiv:2402.17020.
[3]. Injadat, M. N., Salo, F., Bou Nassif, A., Essex, A., & Shami, A. (2020). Bayesian Optimization with Machine Learning Algorithms Towards Anomaly Detection. arXiv preprint arXiv:2008.02327.
[4]. Rahmati, M. (2025). Towards Explainable and Lightweight AI for Real-Time Cyber Threat Hunting in Edge Networks. arXiv preprint arXiv:2504.16118.
[5]. Sewak, M., Sahay, S. K., & Rathore, H. (2022). Deep Reinforcement Learning for Cybersecurity Threat Detection and Protection: A Review. arXiv preprint arXiv:2206.02733.
[6]. Wikipedia contributors. (2025). Adversarial machine learning. In Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Adversarial_machine_learning
[7]. Cypher Scoop. (2024). Leveraging Machine Learning for Anomaly Detection in Cybersecurity. Retrieved from https://www.cypherscoop.com/leveraging-machine-learning-anomaly-detection/
[8]. Journal of Cloud Computing. (2025). AI driven IOMT security framework for advanced malware and ransomware detection in SDN. Retrieved from https://journalofcloudcomputing.springeropen.com/articles/10.1186/s13677-025-00745-w
[9]. Yubetsu Codex. (2024). A Review of Machine Learning Techniques for Anomaly Detection in Cybersecurity. Retrieved from https://codex.yubetsu.com/article/e5c6468c26e84dd5be8829dcd1346f28

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25may2116

[10]. MDPI Algorithms. (2022). AI for Cybersecurity: Robust models for Authentication, Threat and Anomaly Detection. Retrieved from https://www.mdpi.com/journal/algorithms/special_issues/AI_Cybersecurity_Model