

Artificial Intelligence-Driven Cybersecurity: An Adaptive Framework For Network Anomaly And Ddos Detection

Dr Neha Paliwal (Asst. Prof. Computer Science, Shri Mahaveer College, Jaipur)

Abstract: Cybersecurity environments are becoming increasingly dynamic as cloud services, Internet of Things (IoT) devices, mobile platforms and distributed applications generate large and heterogeneous volumes of network activity. Traditional signature-oriented security controls remain valuable for known threats, but they can be less effective when malicious behaviour changes rapidly or resembles legitimate traffic. This study develops an expanded analytical framework for Artificial Intelligence (AI)-enabled cybersecurity, with particular attention to machine learning, deep learning, autoencoders, anomaly detection and Distributed Denial-of-Service (DDoS) detection. The work is grounded in the supplied research article and supplements its analysis with current external evidence from ENISA and NIST and with verified descriptions of widely used intrusion-detection datasets. The original source reports 15 comparative accuracy values ranging from 82.3% to 99.95%; secondary statistical analysis of those values gives a mean of 96.42%, a median of 98.10% and a range of 17.65 percentage points. These figures are treated as reported results, not as newly collected experiments. The study further proposes a human-in-the-loop architecture in which network traffic is captured, cleaned, normalized and represented through an autoencoder before anomaly scoring, classification and analyst verification. Current ENISA evidence indicates that DDoS represented 77% of reported incidents in its 2025 threat landscape covering 4,875 incidents from July 2024 to June 2025. The analysis demonstrates the potential of AI while emphasizing that accuracy alone is insufficient for cybersecurity evaluation. Reliability, false-positive control, explainability, computational cost, concept drift, adversarial robustness and human oversight must also be considered. The article concludes with practical recommendations for multi-dataset validation, continuous monitoring, explainable AI and responsible integration with established cybersecurity risk-management frameworks.

Keywords: Artificial Intelligence; Cybersecurity; Machine Learning; Deep Learning; Autoencoder; DDoS; Intrusion Detection; Anomaly Detection; Explainable AI; Network Security.

1. Introduction

The modern digital environment is characterized by extensive interconnection among users, devices, applications, cloud platforms and communication networks. This connectivity improves accessibility and productivity, but it also creates a large attack surface. Security teams must process high volumes of events while dealing with rapidly changing attack behaviours, distributed infrastructure and limited human attention. The supplied source identifies device proliferation, diverse attack vectors, continuously exploitable vulnerabilities, data overload and shortages of skilled security personnel as major challenges that motivate the use of AI.

AI is attractive in this setting because it can process large numbers of observations and learn statistical relationships from historical or streaming data. Machine-learning classifiers can support traffic categorization, while deep-learning models can learn more complex representations. Autoencoders provide a complementary capability because they can learn compact representations of data and use reconstruction behaviour as an indicator of unusual observations.

The operational importance of DDoS detection is particularly clear in recent threat intelligence. ENISA's 2025 Threat Landscape analysed 4,875 incidents between 1 July 2024 and 30 June 2025 and reported DDoS as the dominant incident type, accounting for 77% of reported incidents. Ransomware was identified as the most impactful threat. This contemporary evidence reinforces the relevance of intelligent availability-threat detection.

However, a high classification score does not automatically establish operational usefulness. Benchmark datasets may differ in traffic composition, attack distribution, feature engineering, labeling and collection methodology. Therefore, this article focuses not only on detection accuracy but also on robustness, generalization, interpretability and deployment considerations.

2. Problem Statement

The central problem is the difficulty of detecting changing network attacks quickly and reliably using methods that depend heavily on previously known signatures. Network traffic can contain noise, redundant features, class imbalance and previously unseen behaviours. An AI system must therefore identify meaningful representations without creating excessive false alarms. The research problem can be expressed as follows:

How can an adaptive AI-based framework combine representation learning, anomaly detection and human verification to improve the identification of DDoS and other network anomalies while remaining interpretable, robust and computationally practical?

3. Research Objectives

- Examine current applications of AI and machine learning in cybersecurity.
- Analyse the role of autoencoders in network-data representation and anomaly detection.
- Compare reported AI/ML detection performance from the supplied source.
- Integrate current threat-landscape evidence into the cybersecurity analysis.
- Propose an adaptive human-in-the-loop architecture for network anomaly detection.
- Identify methodological limitations and provide recommendations for future validation.

4. Research Questions

1. RQ1: What capabilities make AI useful for modern cybersecurity monitoring?
2. RQ2: How does reported performance vary across algorithms and intrusion-detection datasets?
3. RQ3: How can autoencoders support dimensionality reduction and anomaly detection?
4. RQ4: What limitations prevent benchmark accuracy from being treated as proof of real-world effectiveness?
5. RQ5: What architecture can combine AI automation with human cybersecurity expertise?

5. Review of Literature

5.1 Machine Learning for Intrusion Detection

Machine learning has been widely investigated for intrusion detection because network events can be represented as structured feature vectors. Supervised approaches learn relationships between labeled traffic classes, whereas unsupervised and semi-supervised approaches can focus on deviations from normal behaviour. The supplied article reports the use of decision trees, neural networks, Naïve Bayes, ensemble methods, LSSVM, LSTM and AdaBoost across several datasets.

5.2 Deep Learning

Deep learning extends machine learning by using multiple computational layers to learn hierarchical representations. In cybersecurity, this is useful where relationships among packet, flow and behavioural attributes are difficult to capture with manually engineered rules. The trade-off is greater computational demand and potentially reduced interpretability.

5.3 Autoencoders

An autoencoder contains an encoder that maps input data to a latent representation and a decoder that reconstructs the input. When trained primarily on representative normal observations, unusual traffic can produce larger reconstruction errors. The supplied source emphasizes denoising, compression and dimensionality reduction as key benefits.

5.4 DDoS and Availability Threats

DDoS attacks attempt to exhaust resources or otherwise reduce service availability. AI systems can examine flow rates, packet frequencies, source distributions and temporal patterns to identify traffic behaviour that differs from established baselines. Recent ENISA evidence shows that availability threats remain highly prominent.

5.5 Explainability and AI Risk

Cybersecurity analysts require actionable explanations rather than opaque scores. NIST CSF 2.0 provides an organization-wide framework for managing cybersecurity risk, while NIST's AI risk-management work emphasizes trustworthy AI practices. Consequently, AI-based detection should be embedded within governance, monitoring and response processes rather than operated as an isolated prediction engine.

6. Research Gap

Table 1. Research gaps and proposed responses

Gap	Observed limitation	Research response
-----	---------------------	-------------------

Dataset dependence	Performance changes across datasets and attack classes.	Recommend cross-dataset validation.
False positives	Legitimate high-volume activity may resemble attacks.	Use calibrated thresholds and analyst feedback.
Concept drift	Network behaviour changes over time.	Introduce monitoring and periodic model evaluation.
Explainability	Deep models can be difficult to interpret.	Add explanation and feature-attribution layers.
Adversarial robustness	Attackers may manipulate AI inputs or models.	Include robustness testing and defensive monitoring.
Operational cost	Deep models may require substantial resources.	Evaluate latency and compute requirements.
Human dependence	Automated decisions can lack context.	Use human-in-the-loop verification for important alerts.

7. Dataset and Data Sources

The supplied article describes a packet-oriented CSV dataset produced after network capture and feature extraction. The dataset includes packet-level and contextual information and is intended for preprocessing and AI analysis. For expanded contextual analysis, this article also considers established public benchmark datasets.

Table 2. Dataset landscape used for the expanded research design

Dataset	Verified characteristics	Relevance
NSL-KDD	Refined intrusion-detection benchmark derived from KDD-style traffic.	Historical baseline for IDS comparisons.
UNSW-NB15	Generated from a hybrid of real normal activity and synthetic contemporary attacks; 49 features and nine attack types are described by UNSW.	Useful for modern multi-class intrusion analysis.
CIC-IDS2017	Contains benign and common attacks with PCAP and labelled flow data.	Useful for realistic flow-based evaluation.
BoT-IoT	IoT-focused benchmark used in the source comparison.	Relevant to IoT and DDoS analysis.
KDD99	Legacy benchmark represented in several reported results in the supplied source.	Useful for historical comparison, but should not be treated as a complete representation of modern traffic.

UNSW confirms that UNSW-NB15 was created using the IXIA PerfectStorm tool in a cyber range, with 100 GB of raw traffic, nine attack categories and 49 generated features. CIC-IDS2017 provides PCAPs and labelled flows designed for intrusion detection evaluation. These characteristics illustrate why multi-dataset testing is preferable to relying on a single benchmark.

8. Research Methodology

The study follows a structured secondary-analysis and framework-development methodology. It combines (a) synthesis of the supplied article, (b) quantitative analysis of its reported accuracy values, (c) verification of selected dataset characteristics through authoritative dataset sources, (d) current threat-landscape contextualization, and (e) development of a proposed AI detection architecture.

Table 3. Research methodology stages

Stage	Activity	Output
1	Literature and source review	Research themes and gaps
2	Dataset assessment	Dataset characteristics and suitability
3	Data preparation design	Cleaning, encoding, scaling and feature selection

4	Representation learning	Latent feature representation through autoencoder
5	Anomaly analysis	Reconstruction error and threshold-based alerting
6	Comparative analysis	Reported algorithm performance comparison
7	Interpretation	Findings, limitations and deployment implications
8	Framework design	Human-in-the-loop cybersecurity architecture

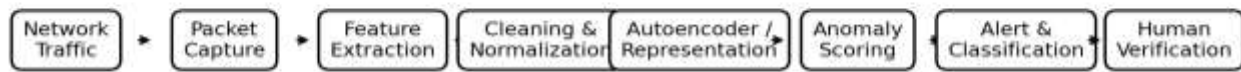


Figure 1. Proposed AI-enabled cybersecurity detection workflow.

9. Data Preprocessing

9.1 Data Cleaning

Remove or document missing values, duplicate observations and malformed records. The exact procedure should be reported because aggressive cleaning can change the distribution of network events.

9.2 Encoding

Categorical protocol, service or state attributes should be transformed into numerical representations appropriate to the selected model.

9.3 Normalization

Continuous features should be scaled so that variables with large numerical ranges do not dominate optimization.

9.4 Feature Selection

Highly redundant or irrelevant variables should be examined. Feature selection can reduce computational cost and improve model stability.

9.5 Training and Testing Separation

Training and testing observations must be separated before model fitting. Where temporal information is available, chronological separation can provide a stronger test of generalization than random splitting.

10. Autoencoder Model

Let x represent an input vector containing normalized network features. The encoder produces a latent representation z :

$$z = f\theta(x)$$

The decoder then reconstructs the input:

$$\hat{x} = g\phi(z)$$

A common reconstruction objective is mean squared error:

$$MSE = (1/n) \sum (x_i - \hat{x}_i)^2$$

If the model learns representative normal traffic, an unusually high reconstruction error can be used as an anomaly signal. The threshold should be determined from validation data rather than selected arbitrarily.

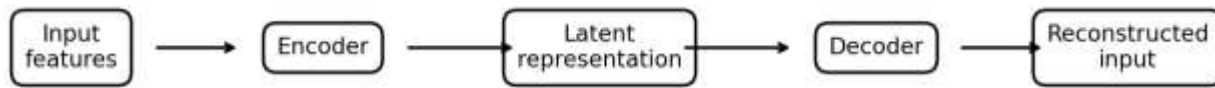


Figure 4. Autoencoder representation-learning structure.

11. Evaluation Metrics

Table 4. Evaluation measures for AI-based cybersecurity

Metric	Formula / interpretation	Why it matters
Accuracy	$(TP+TN)/(TP+TN+FP+FN)$	Overall correctness; can mislead under class imbalance.
Precision	$TP/(TP+FP)$	Measures how many alerts are actually positive.
Recall	$TP/(TP+FN)$	Measures how many attacks are detected.
F1-score	$2PR/(P+R)$	Balances precision and recall.
False Positive Rate	$FP/(FP+TN)$	Important for analyst workload.
False Negative Rate	$FN/(FN+TP)$	Important because missed attacks may be costly.
Detection latency	Time from event to alert	Important for operational response.
Reconstruction error	Distance between input and reconstruction	Core anomaly signal for autoencoders.

12. Quantitative Data Analysis

The supplied article reports 15 comparative accuracy observations. These are treated as secondary data. No claim is made that the present study independently reproduced those experiments.

Table 5. Secondary statistical summary of reported accuracy

Statistic	Value
Number of reported observations	15
Mean	96.42%
Median	98.10%
Minimum	82.30%
Maximum	99.95%
Range	17.65 percentage points
Population standard deviation	5.09 percentage points

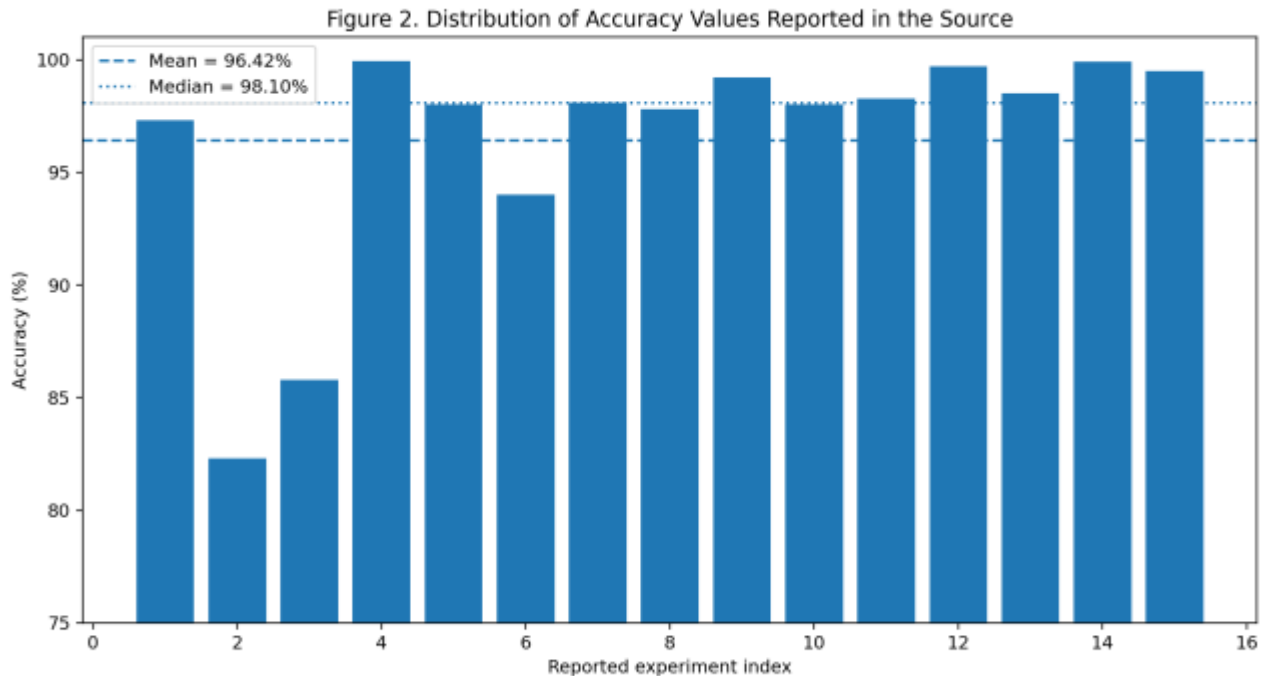


Figure 2. Distribution of accuracy values reported in the supplied source.

13. Comparative Performance Analysis

Table 6. Reported AI/ML results reproduced for secondary analysis

Technique	Dataset	Attack / task	Accuracy
OS-ELM	NSL-KDD	Multiple	97.3%
Neural Network	NSL-KDD	DoS, U2R, R2L	82.3%
Decision Tree + Naïve Bayes	NSL-KDD	Probing, U2R, R2L	85.8%
TAB	KDD99	DoS flooding	99.95%
Decision Tree	KDD99	Multiple attack categories	98.0%
Ensemble Learning	AndroZoo/Drebin	Malware	94.0%
Decision Tree	RPL-NIDDS17	DoS	98.1%
Decision Tree	UNSW-NB15	Multiple attacks	97.8%
Neural Network	NSL-KDD	Probing/U2R/R2L	99.2%
Decision Tree	NSL-KDD	Multiple attacks	98.0%
Neural Network	BoT-IoT	DoS/Reconnaissance/DDoS	98.26%
LSSVM	KDD99	Anomaly detection	99.7%
DFEL	UNSW-NB15	Multiple	98.5%
LSTM	ISCX2012	DoS flooding	99.9%
AdaBoost	UNSW-NB15	Botnet flooding	99.5%

The distribution is concentrated near the upper end of the accuracy scale, but the two lower observations (82.3% and 85.8%) demonstrate why averaging results across different datasets should be interpreted cautiously. A model can achieve very high accuracy on one benchmark while performing less effectively on another because of differences in attack diversity, feature construction and class distributions.

14. Current Threat Context

Table 7. Contemporary cybersecurity context

Indicator	Current evidence	Research implication
ENISA incidents analysed	4,875 incidents, July 2024–June 2025	Large-scale threat monitoring remains necessary.
DDoS share	77% of reported incidents	Availability-threat detection is a major AI application area.
Ransomware	Identified by ENISA as the most impactful threat	AI cybersecurity should not focus on DDoS alone.
Leading intrusion access points	Phishing 60%; vulnerability exploitation 21.3%	Network detection should be complemented by endpoint and identity controls.
NIST CSF 2.0	Framework for managing cybersecurity risk	AI detection should fit within organizational risk management.

Figure 5. DDoS Share of Incidents in ENISA Threat Landscape 2025

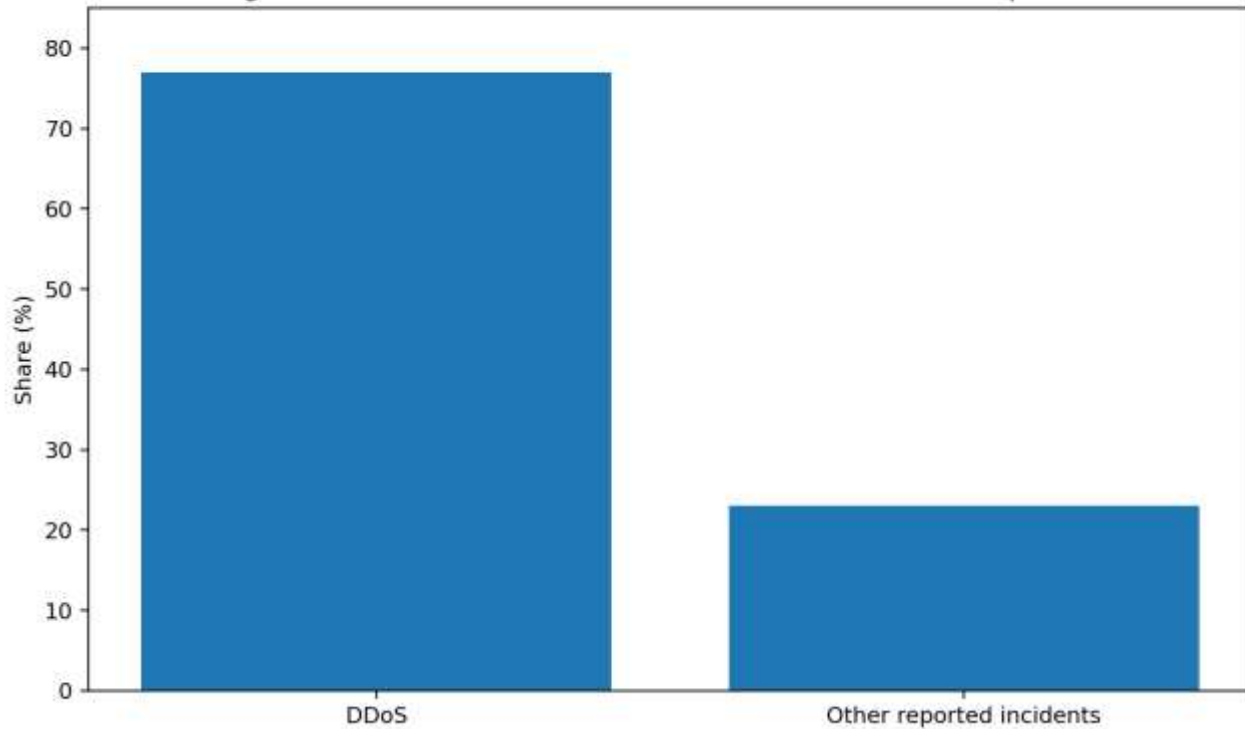


Figure 5. DDoS share of incidents in ENISA Threat Landscape 2025.

15. Dataset Comparison for Future Validation

Table 8. Recommended validation strategy

Dataset	Primary value	Limitation / caution	Recommended use
NSL-KDD	Historical IDS comparison	Older traffic environment	Baseline experiments
UNSW-NB15	Modernized attack categories and 49 features	Cyber-range generation differs from live traffic	Multi-class validation
CIC-IDS2017	PCAP and labelled flow information	Benchmark conditions may not represent every production environment	Flow-based validation
BoT-IoT	IoT-oriented attack scenarios	IoT-specific distribution	IoT/DDoS evaluation
KDD99	Large historical benchmark and comparison baseline	Legacy characteristics	Historical comparison only

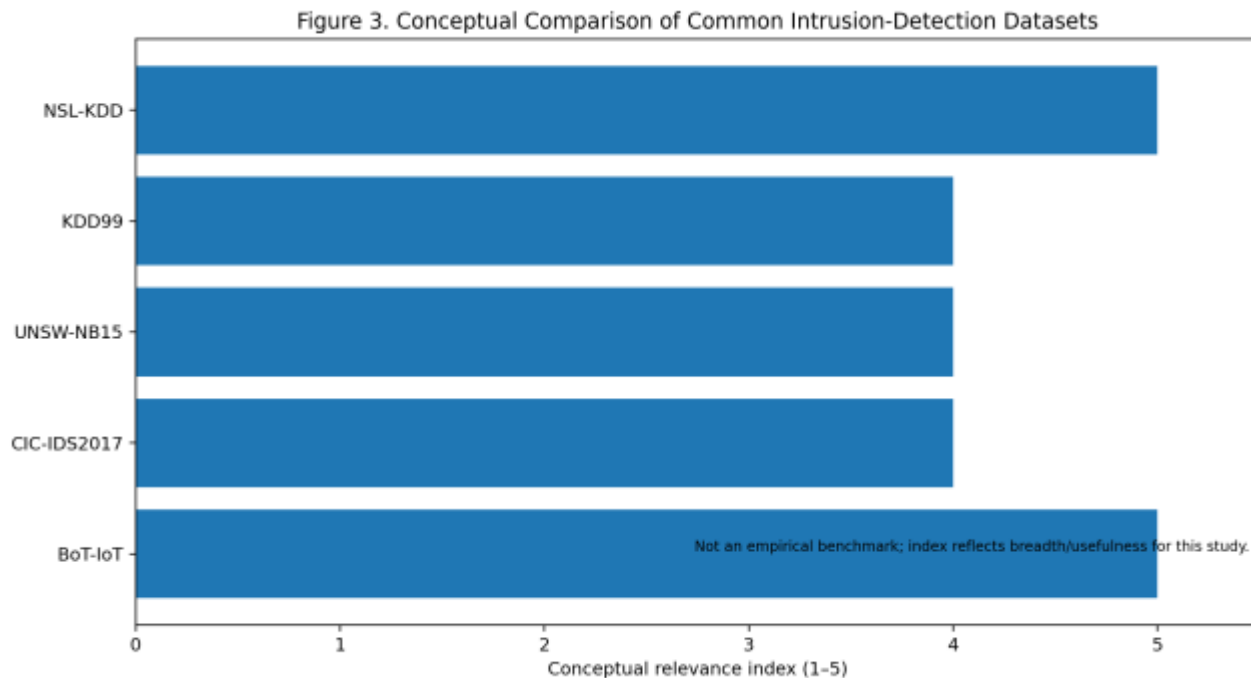


Figure 3. Conceptual comparison of commonly used intrusion-detection datasets. The index is a research-design aid, not an empirical performance score.

16. Findings

Finding 1 — Strong reported AI performance: The secondary dataset contains a mean reported accuracy of 96.42%, indicating substantial potential for AI-assisted detection.

Finding 2 — Performance is context-dependent: The 17.65-point accuracy range shows that performance changes substantially across algorithms, datasets and attack categories.

Finding 3 — Autoencoders are useful beyond classification: Their value lies in representation learning, compression, denoising and anomaly scoring, not merely in assigning attack labels.

Finding 4 — DDoS is a high-priority detection problem: Current ENISA evidence shows DDoS dominating reported incidents in its 2025 threat landscape.

Finding 5 — Accuracy is not enough: False positives, false negatives, latency, robustness and interpretability must be included in operational evaluation.

Finding 6 — Human oversight remains important: Analysts provide contextual judgment that a statistical model may not possess.

Finding 7 — Cross-dataset validation is essential: Results from one benchmark should not be generalized to all network environments without additional testing.

17. Discussion

The evidence supports a balanced view of AI in cybersecurity. AI can substantially improve the ability to process large quantities of network information and identify suspicious patterns. Yet the same properties that make AI powerful—complex representations, automation and adaptation—also introduce new risks. Models can become stale, overfit to benchmark patterns, or produce alerts that analysts cannot easily interpret.

The proposed architecture therefore places the autoencoder within a larger operational loop. Data preparation and representation learning produce an anomaly signal; classification and risk scoring prioritize events; a human analyst reviews important cases; and verified outcomes become feedback for model maintenance. This design is more realistic than assuming that one neural network can independently secure an entire network.

NIST CSF 2.0 is relevant because it frames cybersecurity as a risk-management activity rather than a single technical product. AI detection can therefore support identification, protection, detection, response and recovery activities within a wider organizational program.

18. Proposed Improved Framework

Table 9. Proposed layered AI cybersecurity architecture

Layer	Function	Key controls
Data layer	Collect network and security telemetry	Packet/flow capture, logging, normalization
Representation layer	Learn compact behavioural patterns	Autoencoder, feature learning
Detection layer	Identify unusual or malicious behaviour	Reconstruction error, classifier
Risk layer	Prioritize alerts	Severity, confidence, asset criticality
Human layer	Validate high-impact alerts	Analyst review, investigation
Response layer	Coordinate defensive action	Existing security controls and response processes
Learning layer	Update the model	Feedback, drift monitoring, periodic retraining
Governance layer	Manage AI and cyber risk	Documentation, access control, auditability, policy

19. Practical Suggestions

- Use at least two or three substantially different datasets when evaluating a new detection approach.
- Report precision, recall, F1-score and false-positive rate alongside accuracy.
- Evaluate temporal generalization so that the model is tested on traffic that differs from its training period.
- Use an adaptive threshold for autoencoder reconstruction error and validate the threshold on held-out data.
- Maintain a human-review pathway for high-severity alerts.
- Log model versions, training data versions and threshold changes for auditability.
- Monitor model drift and investigate sudden changes in alert distributions.
- Use explainability techniques to provide analysts with feature-level or event-level evidence.
- Evaluate inference latency and computational requirements before production deployment.
- Integrate AI detection into an established cybersecurity risk-management framework rather than treating it as an independent security control.

20. Limitations

The numerical comparison is a secondary analysis of results reported by the supplied article; it is not a new laboratory benchmark. Because the underlying experiments used different datasets and conditions, direct ranking of algorithms is not methodologically justified. The present work also does not claim a newly measured precision, recall or F1-score because the source does not provide sufficient raw predictions to calculate them independently. The proposed architecture is therefore a research framework requiring future experimental validation.

21. Future Research

- Cross-dataset and cross-environment testing.
- Continual-learning methods for concept drift.
- Federated learning for privacy-preserving collaborative detection.

- Explainable autoencoders and interpretable anomaly scoring.
- Robustness testing against adversarial manipulation.
- Lightweight models for edge and IoT environments.
- Integration of network, endpoint, identity and behavioural signals.
- Real-time evaluation using controlled streaming traffic.
- Standardized reporting of false-positive cost and detection latency.
- Benchmarking against contemporary traffic rather than relying exclusively on legacy datasets.

22. Conclusion

AI offers a significant opportunity to strengthen cybersecurity by transforming large volumes of network observations into actionable security signals. The expanded analysis shows that reported AI/ML studies can achieve high accuracy, but it also shows substantial variation between experiments. The average reported accuracy of 96.42% should therefore be interpreted as a descriptive statistic of the supplied comparison rather than evidence that any single AI model can achieve that level of performance in every environment.

Autoencoders are promising because they provide a mechanism for learning compact representations and identifying unusual network behaviour through reconstruction error. Their most useful role is likely to be within a broader architecture that combines preprocessing, anomaly detection, classification, risk prioritization, human review and continuous learning.

Contemporary threat intelligence reinforces the importance of this direction. ENISA's 2025 landscape reports DDoS as the dominant incident type, while NIST CSF 2.0 provides an organization-wide structure for managing cybersecurity risk. The future of AI-enabled cybersecurity should therefore focus not simply on achieving a higher accuracy number, but on building systems that are reliable, explainable, adaptive, auditable and resilient under changing conditions.

23. References

1. European Union Agency for Cybersecurity (ENISA). (2025). ENISA Threat Landscape 2025.
2. National Institute of Standards and Technology (NIST). (2024). The NIST Cybersecurity Framework (CSF) 2.0. NIST CSWP 29.
3. National Institute of Standards and Technology (NIST). (2024). NIST Cybersecurity Framework 2.0: Resource & Overview Guide. NIST SP 1299.
4. Moustafa, N., & Slay, J. (2015). UNSW-NB15: A comprehensive data set for network intrusion detection systems.
5. Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization.
6. Mirsky, Y., Doitshman, T., Elovici, Y., & Shabtai, A. (2018). Kitsune: An ensemble of autoencoders for online network intrusion detection.
7. Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*, 41(3).
8. Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. *IEEE Symposium on Security and Privacy*.
9. MITRE. (2026). MITRE ATLAS: Adversarial Threat Landscape for Artificial-Intelligence Systems.
10. National Institute of Standards and Technology. (2024). Artificial Intelligence Risk Management Framework resources.