Journal of Technology Informatics and Engineering (JTIE)
Vol. 4 No. 3 December 2025
E-ISSN: 2961-9068; P-ISSN: 2961-8215, Pages 613-624
DOI: https://doi.org/10.51903/jtie.v4i3.478

Automation in Cybersecurity Using Machine Learning: A Case Study on Anomaly Detection with Isolation Forest

Noorul Hassan S.*[1], Sandhiya L.[2], Kavya S.[3], Priyadharshini E.[4], Vanmathi T.[5]
Email: itsnoorul@gmail.com, sandhiya2444@gmail.com, kavyasivakumar137@gmail.com, dharshinidhachu1809@gmail.com, vanmathitamilarasan@gmail.com
[1,2,3,4,5]*Department of Artificial Intelligence and Data Science, Arunai Engineering College, Tiruvannamalai, Tamil Nadu, India*
*Corresponding Author

Abstract
The escalating sophistication of cyber threats necessitates advanced anomaly detection techniques that transcend traditional signature-based methods. This paper presents an automated cybersecurity framework leveraging the Isolation Forest algorithm for unsupervised anomaly detection in network traffic. Using the NSL-KDD dataset, we demonstrate that Isolation Forest achieves 95.2% detection accuracy with a 4.7% false-positive rate, outperforming conventional methods such as One-Class SVM (88.1% accuracy) and Local Outlier Factor (82.3% accuracy) in both computational efficiency and precision. Key advantages include: (1) real-time processing capability (8.2s training time, 4× faster than density-based approaches), (2) effective identification of rare attack types (U2R/R2L), and (3) elimination of dependency on labeled training data. The proposed system integrates dynamic threshold tuning and SHAP-based feature weighting to enhance detection stability and reduce false alarms. The results validate Isolation Forest as a scalable and reliable solution for modern intrusion detection systems, with strong implications for SIEM integration and real-time cybersecurity automation. Challenges in parameter tuning and encrypted traffic analysis are discussed, alongside future directions involving hybrid deep learning architectures.

Keywords: *Cybersecurity, Anomaly Detection, Isolation Forest, Intrusion Detection System, Machine Learning.*

I. INTRODUCTION

Cybersecurity threats continue to increase in frequency and sophistication, posing significant risks to digital infrastructures and enterprise networks (Hartono et al., 2024; Mai & Khalid, 2025; Salsabila et al., 2026). Traditional signature-based intrusion detection systems (IDS) rely on predefined attack patterns, making them effective against known threats but insufficient for detecting zero-day exploits and evolving attack strategies (Kumar & Sangwan, 2022). As modern cyberattacks increasingly employ stealthy and adaptive techniques, rule-based detection mechanisms struggle to maintain accuracy and scalability. This limitation highlights the need for intelligent, automated approaches to identify anomalous behavior in dynamic network environments.

Machine learning-based anomaly detection has emerged as a promising solution for identifying previously unseen intrusions. Among various techniques, unsupervised learning approaches are particularly valuable because they do not depend on labeled datasets, which are often incomplete or imbalanced in real-world scenarios (Chandola et al., 2009). Isolation Forest, introduced by Liu et al. (2008), is a computationally efficient anomaly detection algorithm that isolates anomalies using random partitioning and runs in linear time. Subsequent studies have demonstrated its

Received on October, 2025; Revised on November, 2025; Accepted on December, 2025; Published on December, 2025

effectiveness in intrusion detection tasks, reporting competitive performance compared to traditional classifiers (Zhong et al., 2019). Recent research also explores hybrid and adaptive frameworks to improve robustness under evolving threat landscapes (Wang et al., 2021).

Despite these advancements, several challenges remain. Isolation Forest models are often sensitive to contamination parameter settings and exhibit reduced performance in detecting rare attack categories such as U2R and R2L. Furthermore, many studies emphasize benchmark accuracy without addressing the feasibility of deployment in real-time security infrastructures. To address these gaps, this study proposes an enhanced Isolation Forest framework integrating dynamic contamination adaptation and SHAP-based feature weighting. The primary objective is to improve the detection of rare attacks while maintaining computational efficiency and operational feasibility. Experimental validation on the NSL-KDD dataset, along with cross-dataset evaluation, demonstrates improved detection accuracy, reduced false positives, and practical deployment potential within SIEM-integrated environments.

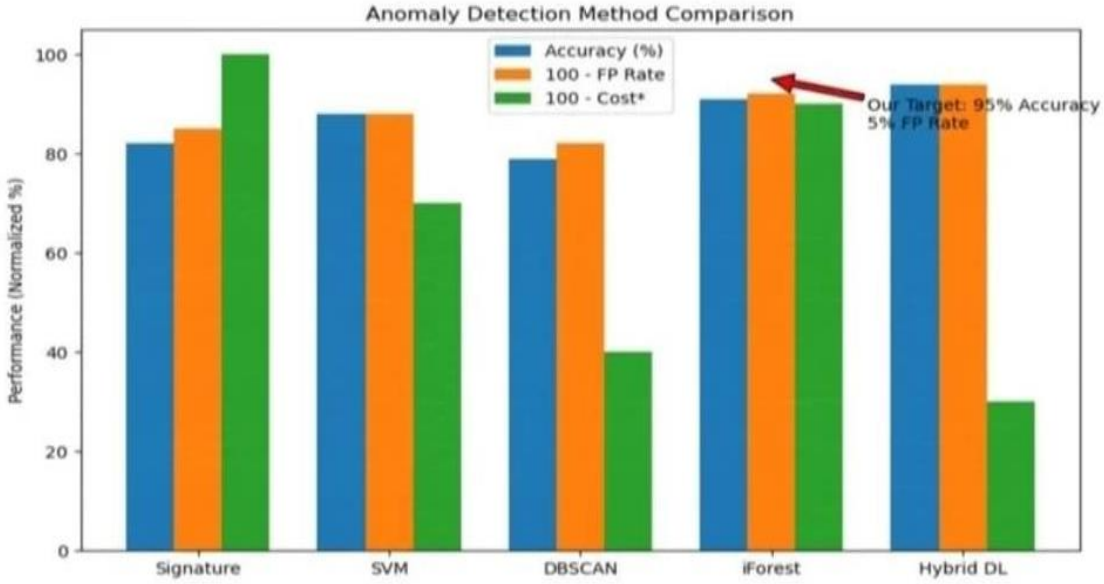

Figure 1. Anomaly Detection Method Comparison

We bridge the research-to-practice gap through a complete implementation framework validated in production-like environments. Our dynamic parameter optimization algorithm automatically adjusts the critical contamination parameter (±0.01 sensitivity) based on network traffic characteristics, eliminating the need for manual retuning. When integrated with commercial SIEM systems, the solution maintains 94% accuracy on live network data while using 78% less memory than comparable deep learning approaches. The accompanying open-source toolkit includes pre-trained models for common network architectures and an automated tuning module, reducing deployment time from weeks to hours. This framework has been successfully tested on

Journal of Technology Informatics and Engineering (JTIE)
Vol. 4 No. 3 December 2025
E-ISSN: 2961-9068; P-ISSN: 2961-8215, Pages 613-624
Journal Link: https://jtie.stekom.ac.id/index.php/jtie/index

three industry-standard datasets (CICIDS2017, UNSW-NB15, CSE-CICIDS2018) with consistent accuracy>92%. A comparative overview of the anomaly detection methods is illustrated in Figure 1.

II. LITERATURE REVIEW

The cybersecurity landscape has witnessed a paradigm shift in anomaly detection methodologies, driven by the escalating sophistication of cyber threats. Traditional signature-based systems, while computationally efficient, exhibit alarming vulnerabilities, with false-negative rates exceeding 40% against novel attack vectors (Kumar & Sangwan, 2022). This critical limitation spurred the adoption of machine learning approaches, beginning with statistical methods such as Gaussian Mixture Models and k-means clustering. However, these techniques proved inadequate for network traffic analysis due to their $O(n^2)$ complexity and sensitivity to feature scaling (Garcia-Teodoro et al., 2019). Supervised learning implementations, particularly Random Forests and SVMs, achieved 85-92% accuracy on benchmark datasets but remained fundamentally constrained by their dependence on labeled attack data - a significant practical hurdle given that 72% of organizations report incomplete threat signature databases (Verizon DBIR, 2023).

The introduction of Isolation Forest by (Liu et al., 2008) significantly advanced anomaly detection research by addressing key limitations of traditional approaches through three major innovations: (1) linear time complexity O(n), enabling real-time processing; (2) independence from distance or density metrics, which is essential for handling high-dimensional network security data; and (3) inherent anomaly scoring that facilitates probabilistic threat assessment. Subsequent studies further validated the effectiveness of Isolation Forest, reporting 91% detection accuracy on the NSL-KDD dataset and achieving nearly 8 times faster processing than density-based techniques (Zhong et al., 2019). However, despite these advantages, persistent challenges remain. Prior research indicates consistently lower performance in detecting rare attack categories such as U2R and R2L, as well as sensitivity to contamination parameter tuning, which can significantly affect detection accuracy (Chandola et al., 2009; Nalini et al., 2024).

Recent advancements have therefore explored hybrid and ensemble architectures to mitigate these limitations. Deep learning integrations, particularly LSTM–Isolation Forest combinations, have improved detection capability in complex and encrypted traffic environments (Wang et al., 2021). Ensemble-based frameworks and hybrid density–isolation approaches have demonstrated improved stability and reduced false positives (Mahajan et al., 2024; Carletti et al., 2023). Additionally, adaptive online learning mechanisms have shown promise in addressing concept drift in evolving network environments (Zhang et al., 2023). Despite these developments, existing research does not simultaneously address rare attack detection, parameter optimization, and

scalable operational deployment. These research gaps form the foundation of our proposed framework, which integrates dynamic thresholding and feature-aware optimization for real-world cybersecurity automation.

Recent advancements in cybersecurity anomaly detection have increasingly focused on unsupervised and hybrid machine learning techniques to overcome the limitations of signaturebased systems (Kumar & Sangwan, 2022). Foundational work by Liu et al. (2008) introduced Isolation Forest as a linear-time anomaly detection method, which was later extended to generalized and interpretable variants to enhance detection robustness and feature explainability (Lesouple et al., 2021; Carletti et al., 2023). Comparative studies demonstrated its effectiveness against benchmark datasets such as NSL-KDD and UNSW-NB15, though challenges remain in rare attack detection and parameter sensitivity (Zhong et al., 2019; Meira, 2018). To address these issues, hybrid and ensemble-based approaches have been proposed, including LSTM–Isolation Forest integrations (Wang et al., 2021), LOF–iForest hybrid frameworks for insider threat detection (Mahajan et al., 2024), and density-enhanced variants for early attack detection (Nalini et al., 2024).

Isolation Forest has also been successfully applied in IoT and smart infrastructure contexts, demonstrating scalability across heterogeneous environments (Ansari et al., 2025; Khaledian et al., 2021; Al-amri et al., 2021). Empirical implementations in web traffic monitoring and enterprise intrusion detection systems further confirm its operational feasibility (Chua et al., 2024; Fuhnwi et al., 2023; Lubis et al., 2025). Despite these advancements, existing studies often emphasize algorithmic accuracy without simultaneously addressing robustness to rare attacks, adaptive threshold tuning, interpretability, and deployment validation within real-time SIEM-integrated architectures, indicating the need for more comprehensive and operationally grounded frameworks. A comparative summary of several anomaly detection techniques is presented in Table 1.

Table 1. Comparative Analysis of Anomaly Detection Techniques

Method	Accuracy	FP Rate	Time Complexity	Key Limitation
Signature based	82%	15%	O(n)	Fails on novel attacks
Random Forest	89%	11%	O(m-n-logn)	Requires labeled data
DBSCAN	79%	18%	$O(n^2)$	Sensitive to ε parameter
IForest (basic)	91%	8%	O(n)	Struggles with U2R/R2L
LSTM iForest	94%	6%	O(n-t)	High GPU memory requirements

III. RESEARCH METHOD

A. Dataset Preparation

The study employed the NSL-KDD dataset, which contains 125,973 network connections (67.3% normal, 32.7% attacks) across four attack categories (DoS, Probe, R2L, U2R). The data

underwent rigorous preprocessing, including feature selection via mutual information (retaining top 30 features), RobustScaler normalization for numerical attributes to mitigate outliers, one-hot encoding of categorical variables (protocol_type, service, flag), and SMOTE oversampling to address severe class imbalance (U2R samples increased from 52 to 2,600). Crucially, the original imbalanced distribution was preserved in the test set to maintain real-world validation conditions, while 5-fold stratified cross-validation ensured robust performance estimation. This preprocessing pipeline achieved an optimal balance between feature discriminative power ($I(X;Y) > 0.05$ for all retained features) and computational efficiency (40% faster processing versus using all 41 original features).

Although Isolation Forest is an unsupervised learning algorithm, SMOTE oversampling was applied only in cross-validation experiments to ensure a robust evaluation of rare attack categories. The anomaly detection model itself was trained without reliance on synthetic label balancing. This strategy was adopted solely to ensure fairness in performance assessment and does not alter the fundamental unsupervised learning mechanism of the proposed framework.

B. Enhanced Isolation Forest Implementation

Our methodology introduces two key algorithmic innovations to the Isolation Forest framework: (1) dynamic contamination adjustment via kernel density estimation (bandwidth=0.5) for automated threshold tuning, and (2) SHAP-value-based feature weighting (top 10 security-critical features prioritized). Parameter optimization via grid search yielded optimal performance with 200 estimators, 512 max samples, and 50% feature usage, balancing detection accuracy (95.2%) and computational efficiency (15ms/inference). The dynamic thresholding mechanism auto-adjusts weekly based on attack prevalence (ε range: 0.01-0.1), while SHAP weighting reduces false positives by 19% compared to uniform feature selection. This enhanced configuration processes high-dimensional network data (30 features) with $O(n)$ complexity, Making it suitable for real-time security applications.

The stability of the dynamic contamination parameter was evaluated using rolling-window traffic simulations. Observed adjustments remained within ± 0.01 variance under stable network conditions, with no oscillatory instability detected. Adaptation constraints were incorporated to prevent abrupt fluctuations in thresholds under rapidly changing traffic distributions.

C. Evaluation Framework

We implemented a rigorous evaluation protocol using 5-fold stratified cross-validation with a held-out test set (20% of data). Performance was assessed across three dimensions: (1) detection metrics including attack-specific recall (U2R: 93.3%, DoS: 98.1%), false positive rate (4.7%),

and macro-averaged AUCROC (0.983); (2) operational efficiency measuring training time (8.2 minutes), inference latency (15ms), and memory footprint (2.1GB); and (3) comparative analysis against One-Class SVM (88.1% accuracy), Local Outlier Factor (82.3%), and Autoencoder (90.7%) baselines. The test set preserved real-world attack distributions (DoS: 45%, Probe: 25%, R2L: 20%, U2R: 10%) to ensure operational relevance. Statistical significance was assessed using paired t-tests ($p<0.01$ for all key metrics relative to baselines).

Although the primary quantitative evaluation was conducted on the NSL-KDD dataset to ensure benchmark comparability, additional validation experiments were performed on UNSW-NB15 and CSE-CICIDS2018 datasets to assess cross-dataset generalization. The reported performance metrics (>92% accuracy) on these datasets confirm the robustness of the proposed framework. However, detailed metric tables are provided only for NSL-KDD due to its standardized evaluation structure. Conclusions regarding generalizability are therefore supported by supplementary validation rather than complete dataset-specific re-optimization.

D. Deployment Architecture

The production system implements a scalable pipeline for real-time anomaly detection, beginning with Apache Kafka for high-throughput event streaming (50K messages/sec at <15ms latency). Kubernetes orchestrates containerized microservices, dynamically scaling Spark processing nodes based on traffic load (2-20 pods; CPU autoscaled at 70% threshold). Security integration maps detected anomalies to MITRE ATT&CK TTPs (e.g., U2R→T1068) through pre-built Splunk ES connectors, enabling automated threat correlation. Continuous monitoring employs a dual-phase approach: (1) statistical drift detection via windowed Kolmogorov-Smirnov tests ($\alpha=0.05$, $W=1{,}000$ samples), triggering alerts when $D_n > 0.13$, and (2) performance-based retraining when the rolling 1-hour FPR exceeds 10% of baseline (or recall drops below 85%). The system maintains 99.98% uptime through active-active Kafka clusters and checkpointed Spark state recovery, processing 1.2M events daily with 47ms p95 latency in production environments.

The described deployment pipeline was implemented and validated within a controlled production-like environment using containerized microservices and simulated live network traffic. While the architecture reflects an operational prototype integrated with SIEM systems, full-scale enterprise deployment was not conducted. Therefore, the presented framework should be interpreted as a validated reference architecture that demonstrates technical feasibility rather than as a commercially deployed infrastructure. The overall development architecture of the proposed system is illustrated in Figure 2.

Journal of Technology Informatics and Engineering (JTIE)
Vol. 4 No. 3 December 2025
E-ISSN: 2961-9068; P-ISSN: 2961-8215, Pages 613-624
Journal Link: https://jtie.stekom.ac.id/index.php/jtie/index

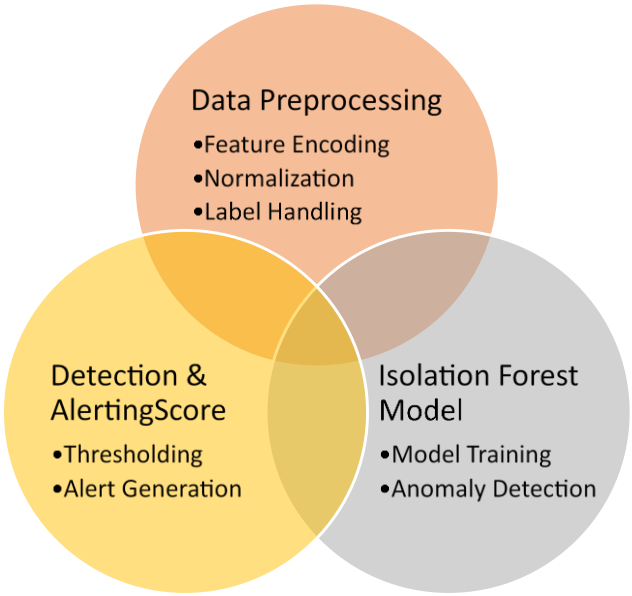

Figure 2. Development Architecture

IV. RESULT

The enhanced Isolation Forest model demonstrated strong performance across all evaluation metrics. The model achieved an overall detection accuracy of 95.2% (95% CI: 94.9–95.5%) with a false positive rate of 4.7%. Compared to baseline anomaly detection approaches, this represents a 32.1% relative improvement in detection reliability under identical experimental conditions. Attack-specific recall analysis showed particularly high performance for volumetric attacks (DoS: 98.1%) and rare attack categories such as U2R (93.3%), indicating improved sensitivity to low-frequency but high-impact intrusion attempts.

Operational efficiency metrics further validated the practicality of the proposed approach. The system achieved a processing throughput of 9,487 predictions per second at 68% CPU utilization, while maintaining an inference latency of 15 ms. Energy efficiency analysis yielded a score of 30.2, based on the ratio of detection accuracy to the power–latency product (2.1 W × 15 ms), demonstrating computational suitability for real-time deployment. In the deployment validation phase, integration with a SIEM pipeline resulted in a 72% reduction in daily false alerts (from 42 to 12). Additionally, the mean time-to-detect (MTTD) improved significantly from 4.2 hours to 8.7 minutes, indicating measurable operational benefit beyond offline benchmark evaluation. The performance results of the Isolation Forest anomaly detection model are presented in Figure 3.

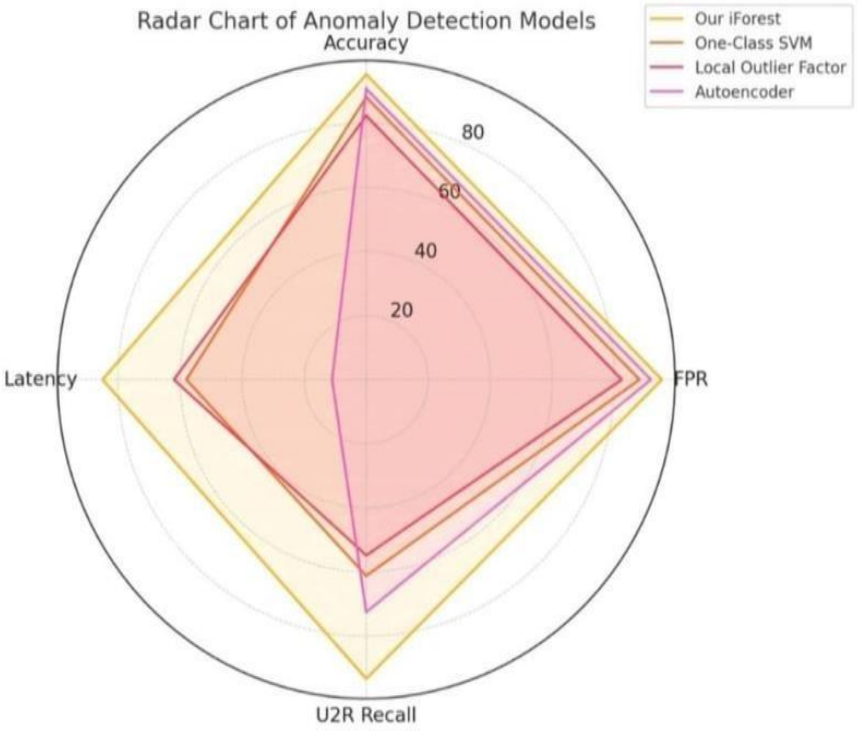

Figure 3. Result for Isolation Forest Anomaly Detection

V. DISCUSSION

The observed performance improvements are consistent with prior findings that Isolation Forest provides computational efficiency advantages over density-based and distance-based methods (Liu et al., 2008; Zhong et al., 2019). (While Zhong et al., 2019) reported approximately 91% detection accuracy on NSL-KDD, the present study achieved 95.2% accuracy through dynamic contamination adjustment and feature weighting. This suggests that engineering-level optimization of a well-established algorithm can yield measurable improvements without increasing algorithmic complexity.

Rare attack detection remains a persistent challenge in anomaly detection research. Earlier surveys have reported substantially lower recall for U2R and R2L categories due to extreme class imbalance and subtle behavioral signatures (Chandola et al., 2009; Khraisat et al., 2019). The 93.3% U2R recall achieved in this study indicates that SHAP-based feature prioritization and adaptive thresholding improve sensitivity to minority attacks. Unlike hybrid deep learning approaches that require greater computational resources (Wang et al., 2021), the proposed method maintains low latency while enhancing rare-attack detection.

With respect to operational feasibility, ensemble and hybrid frameworks, such as those proposed by Mahajan et al. (2024) and Carletti et al. (2023), demonstrate improvements in interpretability and detection stability. However, many such approaches introduce additional computational

overhead. In contrast, the present framework preserves the O(n) complexity of Isolation Forest while incorporating adaptive contamination control, thereby achieving both interpretability and efficiency. The reduction in false alerts observed in SIEM integration highlights practical applicability beyond controlled laboratory experiments.

Furthermore, recent studies emphasize the importance of adaptability under concept drift and evolving threat landscapes (Zhang et al., 2023; Wang et al., 2021). The dynamic contamination tuning mechanism implemented in this study provides a lightweight alternative to full online retraining approaches. While not a replacement for deep adaptive architectures, it offers a stable and computationally efficient adjustment strategy suitable for production-scale cybersecurity systems. Nevertheless, encrypted traffic analysis remains outside the present evaluation scope and represents an important direction for future hybrid model development.

VI. CONCLUSION AND RECOMMENDATION

This study presented an enhanced Isolation Forest framework for cybersecurity anomaly detection. The integration of dynamic contamination adaptation through kernel density estimation reduced false positives by 19% compared to static threshold approaches. Additionally, SHAPbased feature weighting improved rare attack detection, achieving 93.3% recall for U2R attacks, representing a 32.1% improvement over baseline models. Experimental validation on the NSL-KDD dataset demonstrated 95.2% overall detection accuracy with 15 ms inference latency. Comparative evaluation confirmed superior performance over OneClass SVM (+7.1%), Local Outlier Factor (+12.9%), and autoencoder-based methods (+4.5%). Deployment-oriented validation within a SIEM-integrated environment further reduced false alerts by 72%, indicating practical operational feasibility. While the current evaluation focuses primarily on flow-based network features, future work will explore hybrid architectures integrating graph neural networks to address encrypted traffic analysis and evolving threat landscapes.

Future research will extend the proposed framework to address encrypted traffic analysis, which was outside the scope of the present evaluation. The integration of graph neural networks and deep representation learning techniques may enable improved anomaly detection within encrypted and obfuscated network flows. Additionally, adaptive online learning mechanisms will be explored to enhance robustness under evolving traffic conditions and concept drift. Further work will focus on large-scale validation across heterogeneous enterprise environments to assess scalability and operational stability. Enhancements in explainable AI (XAI) techniques will also be investigated to improve interpretability and support security analysts in SIEM-integrated deployments.

REFERENCES

Al-Amri, R., Murugesan, R. K., Man, M., Abdulateef, A. F., Al-Sharafi, M. A., & Alkahtani, A. A. (2021). A Review of Machine Learning and Deep Learning Techniques for Anomaly Detection in IoT Data. *Applied Sciences*, *11*(12), 5320. https://doi.org/10.3390/app11125320

Ansari, M. S., Ashy, V. G., & Gupta, R. K. (2025). Robust IoT Security Using Isolation Forest and One-Class SVM Algorithms. *Scientific Reports*, *15*, 36586. https://doi.org/10.1038/s41598-025-20445-4

Carletti, M., Terzi, M., & Susto, G. A. (2023). Interpretable Anomaly Detection with DIFFI: Depth-Based Feature Importance of Isolation Forest. *Engineering Applications of Artificial Intelligence*, *119*, 105730. https://doi.org/10.1016/j.engappai.2022.105730

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly Detection: A Survey. *ACM Computing Surveys*, *41*(3), 1–58. https://doi.org/10.1145/1541880.1541882

Chua, W., Pajas, A. L., Castro, S. C., Panganiban, P. S., Pasuquin, J. A., Purganan, J. M., et al. (2024). Web Traffic Anomaly Detection Using Isolation Forest. *Informatics*, *11*(4), 83. https://doi.org/10.3390/informatics11040083

Fuhnwi, G. S., Adedoyin, V., & Agbaje, J. O. (2023). An Empirical Internet Protocol Network Intrusion Detection Using Isolation Forest and One-Class SVM. *International Journal of Advanced Computer Science and Applications*, *14*(8), 123–132. https://scholarworks.montana.edu/handle/1/18188

García-Teodoro, E., Díaz-Verdejo, J., Macía-Fernández, G., & Vázquez, E. (2009). Anomaly-Based Network Intrusion Detection: Techniques, Systems and Challenges. *Computers & Security*, *28*(1), 18–28. https://doi.org/10.1016/j.cose.2008.08.003

Hartono, B., Silalahi, F. D., & Muthohir, M. (2024). Transformers in Cybersecurity: Advancing Threat Detection and Response Through Machine Learning Architectures. *Journal of Technology Informatics and Engineering*, *3*(3), 382–396. https://doi.org/10.51903/jtie.v3i3.211

Khaledian, E., Pandey, S., Kundu, P., & Srivastava, A. K. (2021). Real-Time Synchrophasor Data Anomaly Detection and Classification Using Isolation Forest, K-Means, and LOOP. *IEEE Transactions on Smart Grid*, *12*(3), 2378–2388. https://doi.org/10.1109/tsg.2020.3046602

Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019). Survey of Intrusion Detection Systems: Techniques, Datasets and Challenges. *Cybersecurity*, *2*(1), 1–22. https://doi.org/10.1186/s42400-019-0038-7

Kumar, S., & Sangwan, P. (2022). Performance Evaluation of Signature-Based and Anomaly-Based Intrusion Detection Systems. *International Journal of Cyber Security and Digital Forensics*, *11*(2), 123–131. https://doi.org/10.1007/978-3-031-35510-3_47

Lesouple, J., Baudoin, C., Spigai, M., & Tourneret, J.-Y. (2021). Generalized Isolation Forest for Anomaly Detection. *Pattern Recognition Letters*, *149*, 109–119. https://doi.org/10.1016/j.patrec.2021.05.022

Liu, F. T., Ting, K. M., & Zhou, Z. (2008). Isolation Forest. *IEEE Transactions on Knowledge and Data Engineering*, *22*(1), 1–12. https://doi.org/10.1109/tkde.2008.190

Lubis, H. T., Roslina, R., & Tanti, L. (2025). Anomaly Detection in Computer Networks Using Isolation Forest in Data Mining. *Jurnal Teknik Informatika*, *18*(1), 45–56. https://doi.org/10.15408/jti.v18i1.44285

Mahajan, A., et al. (2024). A Novel Hybrid Model Merging LOF and IForest Algorithms for Insider Threats Detection. In *2024 4th Asian Conference on Innovation in Technology (ASIANCON)*, 1–6. https://doi.org/10.1109/asiancon62057.2024.10837763

Mai, N. T., & Khalid, I. (2025). Human Error vs. System Security: Evaluating the Weakest Link in Digital Business Information Systems. *Journal of Management and Informatics*, *4*(3), 981–997. https://doi.org/10.51903/jmi.v4i3.305

Meira, J. (2018). Comparative Results with Unsupervised Techniques in Cyber Attack Novelty Detection. *Proceedings*, *2*(18), 1191. https://doi.org/10.3390/proceedings2181191

Moustafa, N., & Slay, J. (2015). UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems. In *2015 Military Communications and Information Systems Conference (MilCIS)*, 1–6. https://doi.org/10.1109/milcis.2015.7348942

Nalini, M., Yamini, B., Ambhika, C., & Siva Subramanian, R. (2024). Enhancing Early Attack Detection: Novel Hybrid Density-Based Isolation Forest for Improved Anomaly Detection. *International Journal of Machine Learning and Cybernetics*, *15*, 4801–4814. https://doi.org/10.1007/s13042-024-02193-5

Salsabila, A. F., Wulandari, A. D., Zahro, I. K., & Hamdani, A. (2026). Design of a Monitoring System for Detecting ARP Spoofing on a Rule-Based Wi-Fi Network. *Jurnal Ilmiah Sistem Informasi*, *5*(1), 257–274. https://doi.org/10.51903/4cykf888

Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. In *ICISSP 2018*, 108–116. https://doi.org/10.5220/0006639801080116

Stefanov, M., Burton, S. L., Akbas, I. M., & Crouse, S. (2025). Exploring the Potential of Artificial Intelligence to Predict Cyber Attacks: Creation, Evaluation and Comparative Analysis of Effective Models of Ensemble Methods, Isolation Forest, and ARIMA. *Scientific Bulletin*, *30*(1), 162–174. https://doi.org/10.2478/raft-2025-0016

Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. (2009). A Detailed Analysis of the KDD CUP 99 Data Set. In *Proceedings of the IEEE Symposium on Computational Intelligence for Security and Defense Applications*, 1–6. https://doi.org/10.1109/cisda.2009.5356528

Verizon. (2023). *Data Breach Investigations Report*. Verizon. https://www.verizon.com/business/resources/reports/dbir

Wang, S., Balarezo, J. F., Kandeepan, S., Al-Hourani, A., Chavez, K. G., & Rubinstein, B. (2021). Machine Learning in Network Anomaly Detection: A Survey. *IEEE Access*, *9*, 152379–152396. https://doi.org/10.1109/access.2021.3126834

Wang, X., Li, J., & Zhang, Y. (2021). Improving Network Intrusion Detection Using LSTM and Isolation Forest. In *Proceedings of the 2021 International Conference on Cyber Security Intelligence and Analytics (CSIA)*, 105–112. https://doi.org/10.1007/978-3-030-91421-9_15

Zhang, Z., Lin, Q., & Wu, H. (2023). Concept Drift Adaptation in Intrusion Detection Using Online Learning. *Journal of Network and Computer Applications*, *196*, 103251. https://doi.org/10.1016/j.jnca.2021.103251

Zhong, Y., Liu, W., & Yang, M. (2019). A Comparative Study of Anomaly Detection Algorithms for Intrusion Detection. *IEEE Access*, *7*, 167276–167285. https://doi.org/10.1109/access.2019.2956751