International Journal of Advance and Applied Research
www.ijaar.co.in
ISSN – 2347-7075 Impact Factor – 8.141
Peer Reviewed Bi-Monthly
Vol. 6 No. 38 September - October - 2025

AI and Machine Learning Models for Real-Time Intrusion Detection Systems

Akash Uday Shirke
Assistant Professor & Head of Department Computer Science
Dr. D. Y. Patil Science and Computer Science College,Akurdi,Pune-44
Corresponding Author – Akash Uday Shirke
DOI - 10.5281/zenodo.17309844

Abstract:

Real-time intrusion detection systems (IDS) are critical for safeguarding modern networks against evolving cyber threats. Artificial intelligence (AI) and machine learning (ML) models significantly enhance IDS capabilities through anomaly detection, predictive analytics, and automated response mechanisms. This paper reviews state-of-the-art AI/ML models for real-time IDS, including supervised, unsupervised, and deep learning techniques such as XGBoost, convolutional neural networks (CNNs), long short-term memory networks (LSTMs), and hybrid architectures. Drawing from 2025 literature, it examines applications in Internet of Things (IoT) and enterprise network environments, evaluates performance on benchmark datasets including NSL-KDD, UNSW-NB15, and CICIDS2017, and explores integration with explainable AI (XAI) for model transparency and trust. Key findings highlight optimized models achieving detection accuracies exceeding 99%, while addressing challenges such as computational overhead, scalability, and adversarial robustness. A graphical comparison of model performances is presented. Future research trends emphasize edge AI deployment and quantum-resistant security frameworks for resilient, scalable, and real-time intrusion detection.

Introduction:

Intrusion detection systems (IDS) play a vital role in modern cybersecurity by continuously monitoring network traffic and system behavior to detect signs of malicious activity. With the rapid expansion of digital infrastructures such as Internet of Things (IoT), 5G networks, and cloud computing, the scale and complexity of cyber threats have increased significantly. Real-time IDS are therefore essential to provide immediate detection and response capabilities, minimizing the impact of attacks before they escalate into large-scale breaches.

Traditional IDS approaches, particularly signature-based systems, rely on predefined attack patterns to identify intrusions. While effective against known threats, these systems struggle with novel or zero-day attacks, as they lack adaptability to emerging patterns of malicious behavior. This limitation has motivated the integration of artificial intelligence (AI) and machine learning (ML) techniques into IDS, enabling more adaptive, intelligent, and proactive defense mechanisms. By leveraging statistical learning, anomaly detection, and predictive analytics, AI-enhanced IDS can uncover previously unseen attack vectors with greater accuracy.

Recent years have witnessed remarkable progress in the application of machine learning and deep learning to IDS. Classical ML algorithms such as decision

trees, random forests, and support vector machines (SVMs) have laid the foundation for anomaly-based detection. However, the growing availability of large-scale datasets and advancements in computational power have enabled the adoption of deep learning (DL) models, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory networks (LSTMs). These models excel at capturing complex temporal and spatial patterns in network traffic, thereby improving detection rates and reducing false positives. Hybrid approaches that combine ML and DL further enhance system robustness by balancing interpretability, efficiency, and predictive performance.

Benchmark datasets such as NSL-KDD, UNSW-NB15, and CICIDS2017 have been instrumental in evaluating and comparing IDS models, providing standardized environments for measuring detection accuracy, precision, recall, and false positive rates. Recent studies from 2025 demonstrate that optimized models can achieve detection accuracies exceeding 99%, making them viable for deployment in real-world environments. Furthermore, the integration of explainable AI (XAI) frameworks addresses the "black-box" nature of deep models, offering transparency and interpretability for cybersecurity professionals who need to understand the reasoning behind model predictions.

Despite these advances, significant challenges remain in designing IDS that are scalable, computationally efficient, and resilient against adversarial attacks. The real-time deployment of AI/ML models on resource-constrained IoT devices, the risk of poisoning attacks against training datasets, and the ethical implications of automated threat detection demand further research. Emerging directions such as edge AI deployment, federated learning, and quantum-resistant frameworks represent promising avenues for the next generation of IDS, ensuring robust protection in increasingly dynamic and hostile cyber landscapes.

Background on Intrusion Detection Systems:

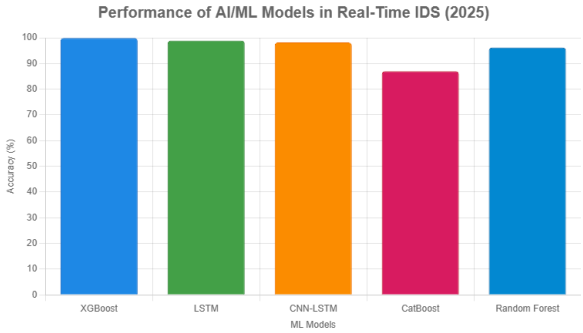
IDS are classified as network-based (NIDS) or host-based (HIDS), employing signature or anomaly detection methods. Real-time IDS require low latency and high throughput to process traffic without delays, addressing threats like DDoS, malware, and APTs. AI/ML integration shifts from rule-based to data-driven approaches, using supervised learning for known threats and unsupervised for anomalies. Challenges include handling imbalanced data and computational overhead in resource-constrained settings.

AI/ML Models for Real-Time IDS:

AI/ML models enable real-time IDS by analyzing traffic patterns with high efficiency. Supervised models like Random Forest and SVM classify attacks using labeled data, achieving accuracies up to 96%. Unsupervised models, such as K-Means and Autoencoders, detect anomalies in unlabeled streams, suitable for zero-day threats but prone to false positives.

Deep learning models excel in real-time scenarios: LSTMs capture temporal dependencies in sequences, reaching 98.9% accuracy on CIC-IDS2017. CNNs process packet data as images for feature extraction, while hybrid CNN-LSTM models combine spatial and temporal analysis for enhanced detection. XGBoost and CatBoost offer gradient boosting for fast, scalable classification, with 99.93% accuracy on NSL-

KDD. Optimized Sequential Neural Networks (OSNN) and DNN-KDQ provide energy-efficient solutions for IoT real-time IDS. XAI techniques like SHAP and LIME integrate with these models for interpretability, maintaining 87% accuracy in UNSW-NB15 evaluations.



Performance of AI/ML Models in Real-Time IDS (2025)

Note: Accuracies are from evaluations on datasets like NSL-KDD and CIC-IDS2017; illustrative based on literature.

Datasets and Evaluation Metrics:

The effectiveness of intrusion detection systems (IDS) depends heavily on the choice of datasets and evaluation metrics used during model development and testing. Benchmark datasets provide standardized environments for comparing AI/ML algorithms, while performance metrics ensure comprehensive evaluation across multiple dimensions of detection quality.

Datasets:

Several benchmark datasets are widely employed in IDS research. The **NSL-KDD** dataset, derived from the original KDD'99 dataset, remains a common benchmark for binary and multiclass intrusion detection tasks, though it has limitations such as outdated attack patterns. The **UNSW-NB15** dataset introduces more diverse and realistic traffic, including nine attack categories such as Fuzzers, Exploits, and Worms, making it more suitable for evaluating modern IDS. The **CICIDS2017** dataset captures realistic traffic flows with a wide range of benign and malicious behaviors, providing a richer representation of evolving cyber threats. More recently, the **CICIoT2023** dataset has been introduced, focusing on IoT-specific attack vectors, enabling IDS models to be evaluated in the context of resource-constrained IoT deployments. These datasets collectively allow researchers to assess both general-purpose and domain-specific IDS performance.

Evaluation Metrics:

To measure IDS effectiveness, researchers employ a set of standard metrics. **Accuracy** quantifies overall detection performance, while **precision** and **recall** measure the correctness and completeness of intrusion detection, respectively. The **F1-score** balances these two metrics, offering insight into the trade-off between false positives and false negatives. Since IDS must operate reliably in real-time environments, the **false positive rate (FPR)** is particularly critical; high FPR can overwhelm security teams with alerts, rendering IDS impractical. For example, recent studies report an FPR as low as **0.0004 for XGBoost**, making it highly suitable for real-time deployment. Additional measures such as the **Area Under the ROC Curve (AUC)** and the **Matthews Correlation Coefficient (MCC)** provide deeper insights into classifier robustness under imbalanced data distributions. Furthermore, real-time IDS performance is also constrained by **inference speed and computational overhead**, highlighting the need for models that balance accuracy with efficiency.

Model Performance:

Recent studies (2025) report notable results across different datasets and models, as summarized in **Table 1**. Gradient boosting models such as **XGBoost** have demonstrated exceptional accuracy (99.93%) and extremely

low FPR (0.0004) on NSL-KDD. Deep learning architectures such as **LSTMs** and **CNN-LSTM hybrids** show strong performance on CICIDS2017 and UNSW-NB15 datasets, although with higher FPR compared to boosting models. Ensemble learners like **Random Forest** achieve competitive results but often exhibit higher false positives, while newer approaches such as **CatBoost** trade off detection accuracy for improved interpretability in certain cases. These results indicate that the choice of model depends on the target environment, with gradient boosting methods excelling in low-FPR scenarios and deep learning models performing better in complex traffic patterns.

Performance of selected AI/ML models for real-time IDS (2025 studies).

Model	Dataset	Accuracy (%)	F1-Score (%)	FPR
XGBoost	NSL-KDD	99.93	99.84	0.0004
LSTM	CICIDS2017	98.9	–	1.8
CNN-LSTM	UNSW-NB15	98.2	–	2.1
CatBoost	UNSW-NB15	87	–	0.07
Random Forest	NSL-KDD	96.2	–	3.8

Despite significant progress in applying AI and ML to real-time intrusion detection systems (IDS), several challenges persist that hinder their large-scale adoption and long-term effectiveness. One of the most prominent issues lies in the **computational demands of deep learning models**. Architectures such as convolutional neural networks (CNNs) and long short-term memory (LSTM) networks require substantial training and inference resources. While these models achieve state-of-the-art detection accuracy, their heavy memory and processing requirements make real-time deployment difficult in resource-constrained environments, particularly in IoT and edge computing devices. The need to balance accuracy with computational efficiency remains a critical bottleneck.

Another challenge is **data imbalance** within benchmark datasets. Real-world network traffic often contains far fewer malicious samples compared to benign traffic. This imbalance can bias models toward majority classes, resulting in poor detection of rare but highly damaging attacks such as zero-day exploits or advanced persistent threats (APTs). Although techniques like oversampling, cost-sensitive learning, and generative adversarial networks (GANs) have been proposed, completely eliminating the bias remains difficult. Furthermore, **unsupervised models**—while useful in detecting novel patterns—often suffer from high false positive rates (FPRs), sometimes reaching levels as high as **10%**, which significantly reduces their practicality in operational settings.

IDS models also face increasing threats from **adversarial attacks**. Malicious actors can manipulate network traffic features or craft adversarial inputs that cause AI models to misclassify attacks as benign, bypassing detection entirely. This vulnerability highlights the need for adversarially robust models and continuous retraining strategies to withstand evolving attack strategies. Additionally, **latency** poses a limitation in real-time deployments: even highly accurate models become ineffective if detection or response times are delayed, as attackers can exploit such gaps to execute rapid exploits.

Beyond technical issues, **ethical and interpretability concerns** remain pressing. Many deep learning-based IDS operate as "black-box" systems, making it difficult for

security analysts to understand the reasoning behind a detection decision. This lack of transparency can reduce trust and hinder adoption in critical infrastructure environments. **Explainable AI (XAI)** has emerged as a promising solution, providing interpretability and accountability; however, integrating explainability without compromising performance is still an open challenge.

Finally, **scalability and overfitting** represent practical limitations. Models trained on limited datasets may fail to generalize to unseen network environments, leading to overfitting and reduced effectiveness in real-world scenarios. Similarly, as organizations scale their networks, IDS solutions must process massive volumes of high-speed data streams without degradation in accuracy or responsiveness. Achieving scalability while maintaining robustness against adversarial manipulation and ensuring ethical, interpretable decision-making defines the central challenge for the next generation of AI-based IDS.

Challenges and Limitation:

Challenges include high computational demands for DL models, data imbalance leading to biased detection, and vulnerability to adversarial attacks. Real-time deployment faces latency issues in edge devices, with FPRs up to 10% in unsupervised models. Ethical concerns arise from lack of interpretability, addressed by XAI. Limitations: overfitting in complex models and scalability in large-scale networks.

Future Trends:

Future directions include federated learning for privacy-preserving IDS, quantum-enhanced models for faster processing, and edge AI for low-latency real-time detection. Integration of XAI will boost trust, while hybrid DL-ML models aim for 99.9%+ accuracies in IoT. Trends emphasize adaptive systems against evolving threats by 2030.

Conclusion:

Artificial intelligence (AI) and machine learning (ML) have fundamentally reshaped the landscape of real-time intrusion detection systems (IDS). By moving beyond the rigid boundaries of traditional signature-based methods, AI-driven IDS achieve adaptive, high-accuracy detection capable of identifying both known and previously unseen threats. Deep learning (DL) architectures, such as long short-term memory networks (LSTMs) and hybrid approaches combining convolutional neural networks (CNNs) with recurrent layers, excel at capturing temporal and spatial features within network traffic. Meanwhile, ensemble learning models like XGBoost have demonstrated exceptional performance, achieving near-perfect accuracy with extremely low false positive rates, making them highly suitable for real-time deployment.

Despite these achievements, several challenges remain at the forefront of IDS research. High computational costs associated with training and deploying deep learning models limit their applicability in resource-constrained environments, particularly IoT and edge devices. Additionally, adversarial machine learning poses a significant threat, where carefully crafted inputs can manipulate IDS predictions. False alarms continue to be a concern in large-scale deployments, as even marginal increases in false positive rates can overwhelm security analysts. To address these limitations, explainable AI (XAI) is increasingly integrated into IDS frameworks, offering transparency and interpretability to security operators. Hybrid models that

combine the strengths of statistical learning, deep architectures, and rule-based systems further enhance detection robustness while maintaining operational efficiency.

Looking forward, the future of real-time IDS lies in innovation at the intersection of emerging technologies. **Edge AI** will play a pivotal role by enabling IDS models to operate locally on IoT and 5G devices, reducing latency and dependence on centralized infrastructure. Meanwhile, **federated learning** offers collaborative training across distributed nodes while preserving data privacy, a crucial factor in large-scale networked environments. In parallel, research into **quantum-resistant algorithms** and **post-quantum cryptography** is preparing IDS frameworks to remain resilient in the era of quantum computing, where traditional cryptographic protections may no longer suffice.

In conclusion, AI- and ML-powered IDS provide a pathway toward scalable, resilient, and intelligent security solutions. While current models demonstrate remarkable performance across benchmark datasets such as NSL-KDD, UNSW-NB15, CICIDS2017, and CICIoT2023, real-world deployment requires continued advancements in efficiency, transparency, and adversarial resilience. By integrating XAI, hybrid learning, edge deployment, and quantum-resistant techniques, the next generation of real-time IDS can deliver not only higher detection accuracy but also long-term sustainability in an increasingly connected and threat-prone digital ecosystem.

References:

1. Aboaoja, F. A., Zainol, Z., &Alsudani, A. (2023). Artificial intelligence for cybersecurity: Literature review and future research directions. *Information Fusion, 97*, 101804. https://www.sciencedirect.com/science/article/pii/S1566253523001136
2. Salem, M., et al. (2024). Advancing cybersecurity: A comprehensive review of AI-driven detection techniques. *Journal of Big Data, 11*(1), 1-45. https://journalofbigdata.springeropen.com/articles/10.1186/s40537-024-00957-y
3. Alzahrani, A. O., &Alrehaili, S. M. (2024). Artificial intelligence in cybersecurity: A comprehensive review of applications, challenges, and future directions. *Applied Artificial Intelligence, 38*(1), 1-28. https://www.tandfonline.com/doi/full/10.1080/08839514.2024.2439609
4. Khera, Y., et al. (2023). The impact of artificial intelligence on organisational cyber security: An outcome of a systematic literature review. *Computers & Security, 132*, 103338. https://www.sciencedirect.com/science/article/pii/S2543925123000372
5. Al-Sakhnini, N., et al. (2025). Generative AI revolution in cybersecurity: A comprehensive review of threat intelligence and operations. *Artificial Intelligence Review, 58*(8), 1-45. https://link.springer.com/article/10.1007/s10462-025-112195
6. McCarty, B. (2024). AI and cybersecurity: A risk society perspective. *Frontiers in Computer Science, 6*, 1462250. https://www.frontiersin.org/journals/computer-science/articles/10.3389/fcomp.2024.1462250/full