**International Journal of Scientific Research and Management (IJSRM)**
**||Volume||13||Issue||07||Pages||2446-2456||2025||**
**Website: https://ijsrm.net ISSN (e): 2321-3418**
**DOI: 10.18535/ijsrm/v13i07.ec07**

# Machine Learning Methods for Intrusion Detection: A Comprehensive Survey

**Mobolaji Ogunbadejo[1*], Oluwatobi Adedamola Ayilara-Adewale[2], Olanrewaju Alade[3]**

[1*]Department of Information System Management, Stanton University, 888 Disneyland Dr #400, Anaheim, California 92802, USA.

[2]Department of Information Technology, Osun State University, P.M.B. 4494, Oke Baale Road, Osogbo, Nigeria.

[3]Department of Information System Management, Stanton University, 888 Disneyland Dr #400, Anaheim, California 92802, USA.

Corresponding Author: **Mobolaji Ogunbadejo[1*]**

**Abstract**
The exponential growth in connected networks, driven by the proliferation of the Internet of Things (IoT) and cloud computing, has resulted in surge in cyberattacks. Advanced and highly sophisticated threats have increased in prevalence, now encompassing advanced persistent threats, distributed denial-of-service attacks, and ransomware. Unfortunately, the signature- and rule-based detection mechanisms used in conventional Intrusion Detection Systems (IDSs) are failing to keep pace, especially with the increasing number of zero-day and newly discovered threats. Machine learning promises to be a futuristic technology due to its capability to identify patterns of activity, autonomously detect new attack designs, and instantly detect deviations in real-time. This survey comprehensively explores and examines the application of supervised, unsupervised, semi-supervised, hybrid, and deep learning methods in Intrusion Detection Systems (IDS), highlighting their unique contributions, strengths, and limitations.

**Keywords:** Intrusion detection, machine learning, anomaly detection, deep learning, cybersecurity.

## 1. Introduction

The digital revolution in contemporary society has led to a groundbreaking increase in network traffic, propelled by a surge in connected devices, the Internet of Things (IoT), and cloud computing. Although this connectivity offers immense benefits, it also provides fertile ground for cybercriminals to exploit vulnerabilities. Sophisticated cyberattacks, including phishing, advanced persistent threats (APTs), Distributed Denial-of-Service (DDoS) attacks, and ransomware, are becoming increasingly prevalent and damaging (Ogunbadejo et al., 2025a). The universal cost of cybercrime is forecasted to exceed $10.5 trillion annually by 2025, emphasizing the urgency for a robust cybersecurity strategy (Sharif and Mohammed, 2022, Akshaya and Saravanan, 2024)

An intrusion detection system (IDS) is a significant approach for protecting network infrastructures by monitoring traffic, identifying malicious actions, and providing prompt alert responses to potential threats, as shown in Figure 1. The conventional IDS employs signature-based detection and rule-based techniques, which are proficient at detecting known attacks but struggle to detect zero-day threats (Zukaib et al., 2024, Rai et al., 2025). Furthermore, the evolving and increasing cyber threats, coupled with the vast scale of network traffic, make it difficult for rule-based systems to keep track of dynamic threats due to their struggle with significant false-positive rates, inability to analyze large amounts of data in real time, and limited adaptability.

Machine learning (ML) has emerged as an innovative technology that enhances Intrusion Detection Systems (IDS) in pattern recognition, automated detection, and real-time anomaly detection (Afridi, 2024). ML helps IDS learn patterns from the past to detect anomalies and understand new forms of attack. Compared to

traditional methods, ML-based IDS can process large amounts of high-dimensional data in real time (Dong and Kotenko, 2025). The capabilities of ML-based IDS have made them efficient for contemporary network environments.

Over the past decade, numerous researchers have investigated a wide range of machine learning (ML) techniques, including traditional algorithms such as Support Vector Machines (SVM) and Decision Trees, as well as deep learning models like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). These techniques are very successful at improving detection accuracy, reducing the number of false positives, and recognizing even the latest threats (Khan and Ghafoor, 2024).

This article presents an in-depth investigation of machine learning methodologies in intrusion detection, laying the groundwork for a more comprehensive study of their applications, current usage, advantages and disadvantages, and shifts in the cybersecurity defense landscape.

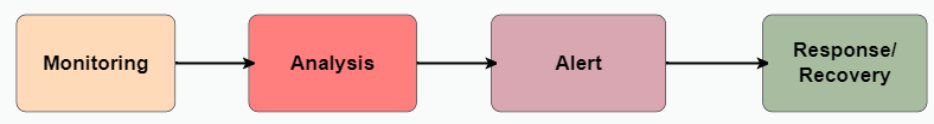


Figure 1: Intrusion detection process

## 2. Overview Of Intrusion Detection Systems

Intrusion detection systems have been categorized into deployment-based and detection approaches, as depicted in Figure 2. The respective category offers unique techniques and applications tailored to each security layer for a specific environment.

### 2.1 Deployment-based intrusion detection systems (IDS)

They are categorized based on their operation within a system; two prominent techniques are Network-Based Intrusion Detection Systems (NIDS) and Host-Based Intrusion Detection Systems (HIDS).

2.1.1 Network Intrusion Detection System (NIDS)

This is a real-time detection system that monitors network traffic by capturing and inspecting packets as they enter the network. These are often installed in strategic locations, including gateways, routers, and switches, where they can evaluate traffic rates on a broad basis and identify indications of an attack (e.g., port scanning, DoS, or potential malware). While NIDS can help detect network-level threats, they may struggle with encrypted traffic and attacks that originate from within the network. Snort and Suricata are two of the most widely used Network Intrusion Detection Systems (NIDS) in the market.

2.1.2 Host-based IDS (HIDS) work by monitoring and analyzing the internals of individual devices or hosts. They use logs, file integrity monitoring, running processes, and sometimes, tracing user behaviors to detect abnormal actions. The merit of HIDS is the ability to precisely identify insider threats and unauthorized access attempts on specific systems. On the other hand, the demerit is that HIDS must be used in every critical host, which may result in higher deployment efforts, increased resource consumption, and the management of all devices. Examples of HIDS tools include OSSEC and Tripwire.

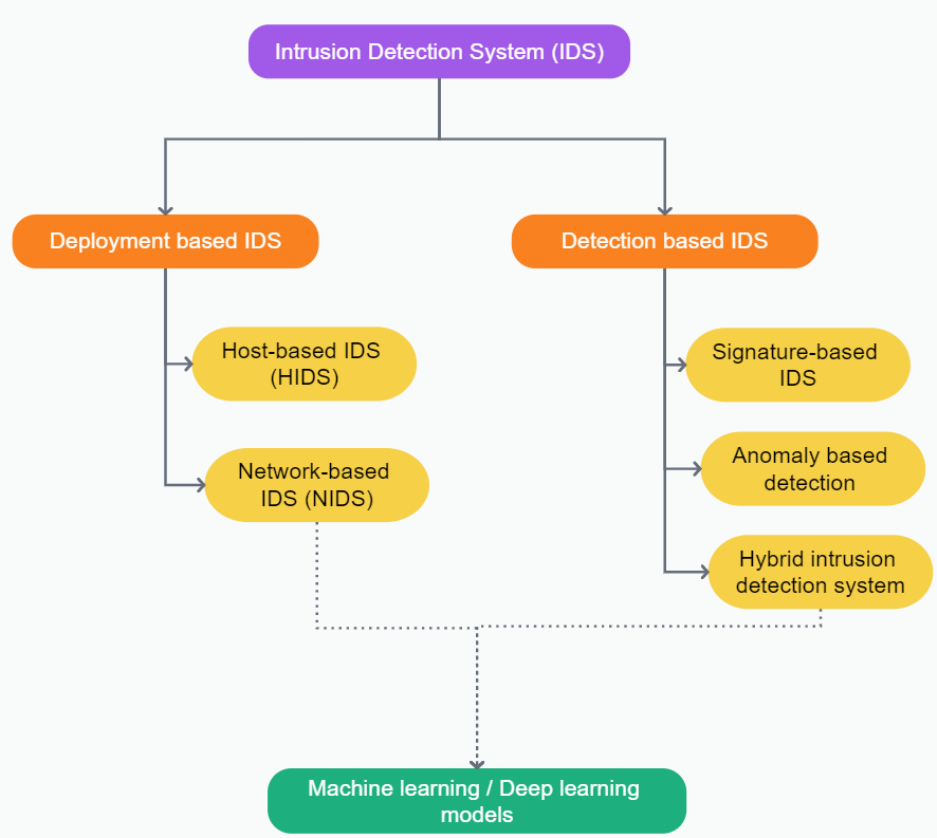

Figure 2: Categorization of Intrusion Detection Systems

## 2.2 Detection Approach-based Intrusion Detection Systems

Intrusion Detection Systems rely on various detection techniques to identify unauthorized and malicious activities in a network or host system. The primary methods are signature-based detection, anomaly-based detection and hybrid intrusion detection systems.

### 2.2.1 Signature-based Detection

In this modality, an IDS matches the behavior it monitors against a database of known attack patterns or signatures. These signatures are functionally predefined if-then rules or patterns activated by recognized threat types such as malware, exploits, or unauthorized access attempts. In practice, the analogy can be drawn with antivirus software, which is very good at identifying and blocking threats it is aware of, yet helpless against whatever modifications or sudden changes in attack strategies, unknown until the next signature update.

### 2.2.2 Anomaly-based Detection.

Anomaly-based IDS models the regular action of a system or network and then labels any deviation from it as potentially malicious threats. Anomaly-based IDS uses statistical analysis, heuristics, or machine learning, implying training on historical data. In this regard, it thrives in environments with high demands for adaptability and early warnings. At the same time, the overall success remains highly contingent on the nature, volume, and quality of the training data, as well as the sensitivity of the anomaly model.

### 2.2.3 Hybrid Intrusion Detection System

To optimize and benefit from both options, many contemporary IDS use multiple methodologies, combining signature and anomaly-based technologies. In effect, this fusion yields accurate and comprehensive detection capabilities, leveraging the precision of signature-based protocols and intelligence in adapting to anomaly models.

## 3. MACHINE LEARNING TECHNIQUES FOR INTRUSION DETECTION

Machine learning (ML) is a key element of anomaly-based intrusion detection systems (IDS), as their mode of operation shares the same principle of learning normal patterns of data and alerting on deviations, i.e., anomalies that could be caused by malicious activity. ML-based anomaly detection mechanisms can be trained to understand "normal" system or network behavior based on the patterns present in the training dataset, as depicted in Figure 3.

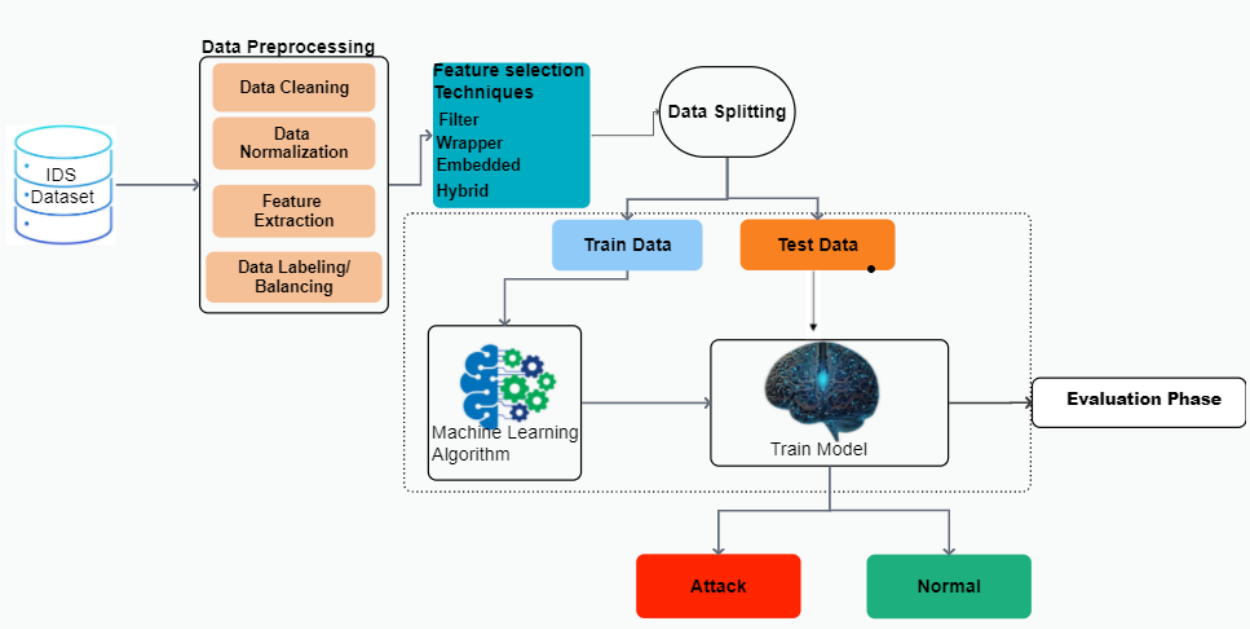


Figure 3: Machine learning-based Intrusion Detection System

As a result, these models are characterized by a rapid adaptation period, which is beneficial for an IDS due to the detection of new threats, such as zero-day attacks, insider threats, and unknown malware (Ogunbadejo et al., 2025b). According to the nature of the data and learning objectives, ML techniques in IDS are categorized into several classes, including supervised, unsupervised, semi-supervised, hybrid, and deep learning. Each class contributes to the efficiency and adaptability of the current IDS in a unique way.

### 3.1 Intrusion Detection Systems Based on Supervised Learning

Supervised learning employs labelled training data to determine the output-label mapping (e.g., normal or malicious); thus, the model learns to predict the sought targets through training on labelled examples. Some of the supervised learning algorithms include decision trees, support vector machines, random forests, K-nearest Neighbours, and Naïve Bayes. Supervised models are efficient at recognizing known attack classes based on previously seen examples. The challenge with this is that it is almost impossible to collect labelled data in the real world due to a combination of the inherent complexity of the labelling process and the existence of new and unknown attack classes.

### 3.2 Intrusion Detection Systems Based on Unsupervised Learning

Unsupervised learning is used when there is no labelled data to train the model. The purpose of these models is to learn patterns, structures, or anomalies in the data without prior knowledge. Several algorithms have been adopted, including K-means clustering, Autoencoder, and principal component analysis (PCA). Similarly, Unsupervised techniques are suitable for situations where new patterns evolve continuously, and zero-day attack detection is required.

### 3.3 Semi-supervised and Hybrid Models

Semi-supervised learning uses a small number of instances labelled by an expert and a large number of unlabeled samples to identify such instances from the top of the list generated by unsupervised methods. This method reduces the need for labels, making semi-supervised ML perfect for ID when labelling instances is costly or impossible. Semi-supervised Models combine labelled data to infer labels of unlabeled examples, using the available instances more efficiently, such as through self-training and label propagation.
Hybrid Models combine several learning strategies and approaches, such as machine learning (ML) and signature-based methods, as well as supervised and unsupervised approaches, to increase the number of intrusions detected and reduce the number of false positives (FP) and false negatives (FN) alerts. Hybrid and semi-supervised models appear to be most useful in real-time systems due to increased scalability, adaptability, and fault tolerance.

### 3.4 Deep Learning Methods

Deep learning methods can automatically learn complex representations from raw data and identify intricate patterns without the need for manual feature engineering. Deep learning models have the following features: an extensive database is required for training, high computational resources are needed, and interpreting the results can be challenging (Taye, 2023). Common types of deep learning methods used in IDS are convolutional neural networks (CNN), recurrent neural networks (RNN) and generative adversarial networks (GAN).

The machine learning techniques used in intrusion detection systems are summarized in Table 1. The adoption of machine learning techniques into IDS has revolutionized the way systems detect and respond to threats. Supervised techniques achieve high accuracy by utilizing known patterns, while unsupervised methods demonstrate excellent performance in detecting new threats. Hybrid and deep learning approaches provided scalability and the ability to adapt. Therefore, it is essential to choose a technology based on the availability of data and computational resources and to consider the security measures adopted in the environment where it will be deployed.

**Table 1: Summary of Machine Learning Algorithms in IDS**

| S/N | IDS Learning Types | Algorithm | Key Features | Suitability |
|---|---|---|---|---|
| 1 | Supervised | Support vector machine (SVM) | Effective in high dimensions | Attack detection, malware classification, |
| | | Random Forest | Robust, ensemble | Large-scale intrusion detection |
| | | Decision Tree | Overfitting-prone, interpretable rules | small to mid-size data, basic classification |
| | | K Nearest Neighbour (KNN) | No training phase, lazy learning, | Lightweight IDS with low data volume |
| | | Naïve Bayes | Probabilistic, fast | Phishing detection / spam |
| 2 | Unsupervised | Principal Component Analysis (PCA) | Visualization, linear transformation, | Outlier detection, feature reduction |
| | | K-Means | Distance-based clustering | Anomaly detection |
| | | Autoencoders | Neural representation learning | Unknown pattern recognition |
| 3 | Semi-supervised | Self-Training | Bootstrapping unlabeled data | Label-scarce environments |
| 4 | Hybrid | Machine learning + signature | Combines behavior and pattern matching | Real-time intrusion detection system |
| 5 | Deep learning | GAN, RNN, CNN | Deep, non-linear pattern discovery | Complex network behavior modelling |

## 3. Literature Survey

The advancements in machine learning-based IDSs over the last three years, as reviewed in this study, specifically cover the works between 2023 and 2025. This literature review is based on recent advancements in intrusion detection systems. The realm of cyber threats has rapidly progressed, necessitating more

innovative and complex detection strategies. Machine learning and deep learning have emerged as powerful tools for identifying network intrusions.

(Mohammed, 2025) proposed an AI intrusion detection system which adopted multiple ML techniques, including Random Forests and SVM, to improve the accuracy of detection while minimizing false positive alerts. The study employed SHAP-based analysis for feature selection of significant attributes in network traffic, aiming to enhance the model's interpretability and performance. The incorporation of reinforcement learning made the response system adaptive and adjustable to new types of threats and heterogeneous network environments.

(Ahmed et al., 2025) identified and evaluated the most efficient machine learning algorithms for IDS systematically. These algorithms include SVM, K-Nearest Neighbors, Random Forest (RF), and Decision Tree. The study's comparative analysis revealed that traditional machine-learning models were effective in coordinating network traffic data and distinguishing between normal and intrusive patterns. Therefore, the study recommends SVM and RF as suitable for real-world IDS applications due to their interpretability, versatility, and reliability as security solutions.

(Dash et al., 2025) introduced an optimized Long Short-Term Memory (LSTM) model for the detection of anomalies in network traffic. The study leveraged the temporal pattern recognition capabilities of LSTM networks to detect sequential patterns of attack that may evade classical detection approaches. The integration of attention mechanisms and hyperparameter optimization enabled the model to display improved capability in detecting complex attack sequences.

(Feng, 2024) presented a machine learning intrusion detection system that employs deep learning via multi-layer perceptron. PyTorch was adopted, which integrated multiple hidden layers to effectively discover non-linear relationships and complex patterns in network traffic data. Furthermore, the study adjusted the learning rate adaptively and introduced an L2 regularization facility to vastly improve generalization. The result is that MLIDS demonstrates a detection accuracy of 98.76%, outperforming traditional techniques such as Naïve Bayes and Single-Layer Perceptron's and revolutionizing malware detection using deep learning techniques.

(Wang et al., 2023a) introduced an unsupervised machine learning algorithm for network detection and defense employing a clustering algorithm. The experimental results showed that the proposed network detection and defense based on an unsupervised machine learning method were significantly better in terms of accuracy and efficiency. The accuracy of traditional methods was 68.49%, whereas the developed system achieved an accuracy of 78.69%. The comparison of the conventional method to the proposed technique in this study reveals that it is less susceptible to unknown network attacks, with reduced false-positive and false-negative rates.

(Wang et al., 2023b) presented the FedVB system, a federated multi-branch neural network intrusion detection system in vertical blocking aggregation. The study aims to maintain privacy in grounded cloud–fog–edge computing environments and develops the ability to train models using the same devices but without direct access to source data. The combination of specialized binary classifiers into a unified architecture enables the detection of specific attack types in anomalous network traffic, providing implementable insights for security components. The result showed a 15.0% higher attack-type matching ratio than conventional multi-class mechanisms, specifically in detecting port scanning attacks and distributed denial-of-service (DDoS) attacks.

(Awajan, 2023) proposed a connected four-layer deep learning architecture for intrusion detection in IoT networks to identify malicious traffic, focusing on connected IoT systems with no reliance on network protocols. The study employed a deep, fully connected (FC) network to identify various attacks, including distributed denial-of-service (DDoS), Sinkhole, Blackhole, and Workhole attacks, as well as opportunistic service attacks. The developed model demonstrates a dependable detection performance of 93.47% in simulated and real-life intrusion scenarios. Additionally, it maintained an average detection rate of 93.21%, which proves its effectiveness in improving IoT network security while optimizing deployment complexity due to its protocol independence.

## 4. Datasets Utilised For Designing IDS

The performance of Intrusion Detection Systems relies mainly on the quality and relevance of the dataset used during their development and testing (Yadav et al., 2020, Tripathy and Behera, 2024).

The efficiency of an IDS relies on the reliability, integrity, and relevance of the data employed during the training and evaluation of the system, as well as the security and authenticity assured by the IDS solutions. Given this, researchers and developers depend on curated datasets and publicly available repositories that can be used to simulate real-world attack events and baseline system activity.

### 4.1 KDD Cup 1999 Dataset

The KDD Cup 1999 dataset is one of the earliest and most widely used benchmark datasets for IDS performance evaluation (Maseer et al., 2021, Thakkar and Lohiya, 2022). It was created for the Third International Knowledge Discovery and Data Mining Tools Competition, held during the 1999 ACM SIGKDD Conference. Today, the KDD Cup 99 is based on the DARPA 1998 intrusion detection evaluation program developed by the MIT Lincoln Laboratory. It is intended to simulate various types of cyberattacks in a military network environment. The KDD dataset contains network traffic data captured over a simulated seven-week period. The respective record in the dataset denotes a single network connection characterized by 41 parameters labelled with different attributes, such as content, traffic and basic features. Similarly, the respective connection is categorized into four key groups: Denial of Service (DoS), Remote to Local (R2L), Probing and User to Root (U2R).

### 4.2 NSL-KDD Dataset

The NSL-KDD dataset was developed as an improved version of the KDD Cup 1999 dataset (Kavitha and Uma Maheswari, 2021). It was created by the University of New Brunswick, Canada, to address and eliminate several flaws of the original KDD'99 dataset. The KDD dataset's deficiencies include an excessive number of duplicate records, a biased evaluation algorithm, and class imbalance (Sahli, 2022). In addition, the NSL-KDD comprises two main sets: KDDTrain+, which consists of clean datasets with 125,973 records, and KDDTEST+, containing 22,544 records (Eshak Magdy et al., 2023, Lin et al., 2024).

NSL-KDD follows the predecessor's structure, which also contains records of 41 features concerning an individual network connection; these tracks are divided into basic, content-based, and traffic-based categories. All connections are assigned to one of the pre-established categories; in the present dataset, there are normal connections and several types of attacks. They are also categorized into four families based on the potential damage they can inflict: Denial of Service (DoS), User to Root (U2R), Remote to Local (R2L), and Probing.

### 4.3 UNSW-NB15 Dataset

UNSW-NB15 is a contemporary benchmark dataset specifically designed to evaluate intrusion detection systems (IDSs) (Disha and Waheed, 2021). It was developed by the Australian Centre for Cyber Security (ACCS) at the University of New South Wales. This was done since previously used datasets such as KDD'99 and NSL-KDD, although widely employed, have certain limitations related to the presence of obsolete attack types and the use of emulated environments, which do not correspond to the modern threat landscape.

The primary objective of UNSW-NB15 was to provide an updated and contextual dataset that describes modern attack behaviors and benign traffic, utilizing contemporary protocols. It represents real-world network behavior by generating traffic through IVIA PerfectStorm, a cutting-edge network traffic generator. Additionally, UNSW-NB15 encompasses various types of attacks, categorized into a broad spectrum of contemporary cyber threats, including backdoors, denial-of-service attacks, exploits, fuzzers, generic attacks, reconnaissance, shellcode analysis, and worms.

### 4.4 TON_IoT Dataset

Telemetry, Operating Systems and Network IoT (TON_IoT) dataset is a benchmark dataset established to advance intrusion detection systems in Internet of Things (IoT) and Industrial IoT (IIoT) environments (Moustafa, 2021). The dataset was created by the Cyber Range Lab of the University of New South Wales (UNSW), Australia. The TON_IoT dataset addresses the evolving need for cybersecurity solutions tailored to innovative ecosystems where conventional datasets are insufficient to capture the sophistication of IoT devices.

TON_IoT was developed to help practitioners and researchers access realistic, comprehensive, and current datasets containing not only network traffic but also telemetry data from IoT sensors, application layer logs

from cloud services, edge devices, and system logs from various platforms, including Linux, Android, and Windows. The dataset comprises numerous modern attack scenarios that mimic real-world threats, including Man-in-the-Middle attacks, password brute-force attacks, ransomware, Botnet activities, DoS/DDoS attacks, backdoor attacks, and data exfiltration.

## 5. Performance Evaluation Metrics

Several key evaluation metrics can be adopted to measure the performance of machine learning models in IDS, and each highlights a different aspect of the model's efficiency. A balanced assessment based on these metrics is vital; focusing solely on one of them can lead to the creation of a model that generates too many false positives or fails to recognize subtle threats. Therefore, to ensure an effective and robust IDS, a comprehensive and context-based evaluation model is required.

### 5.1. Accuracy

Accuracy measures the overall proportion of precisely classified instances, including true positives (TP) and true negatives (TN), in proportion to the total predictions made**.** The equation is:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

In the context of IDS, TP represents the correctly detected intrusions, while TN represents the correctly identified normal traffic. FP represents the false positives (normal traffic incorrectly classified as an intrusion), and FN represents the false negatives (intrusions incorrectly classified as normal traffic).

### 5.2. Precision

Precision can calculate the amount of true positive detections over the number of instances the model predicted as positive. This demonstrates the model's reliability in accurately flagging intrusions, as it indicates how well it avoids false positives (incorrectly labelling benign activity as malware). The equation is:

$$Precision = \frac{TP}{TP + FP}$$

Where TP is True Positives (correctly identified intrusions), and FN is False Negatives (intrusions incorrectly classified as benign), similarly, when the recall value is high, the model is efficient at capturing as many actual intrusions as possible. However, it can introduce false alarms if the precision is low.

### 5.3 Recall

This is also known as the Sensitivity or True Positive Rate, which quantifies the actual positive instances (intrusions) that the model can identify correctly. Similarly, it also demonstrates the model's ability to detect intrusions correctly, making it a good performance metric for an IDS. The equation to calculate this is given as:

$$Recall = \frac{TP}{TP + FN}$$

TP is the True Positives of correctly identified intrusions, while FN is the False Negatives of intrusions incorrectly categorized as benign. High recall means the model can effectively capture the actual intrusion as much as possible; however, if precision is low, it may have introduced false alarms.

### 5.4 F1-Score

The F1-score is the harmonic mean of Recall and Precision, offering a balanced measure of the model's efficiency, particularly in the trade-off between recall and precision. It is an essential IDS metric because false-positive and false-negative errors can affect the system's effectiveness. A high F1 score indicates that the model achieves a good balance between correctly identifying intrusions and minimizing false alarms. The equation for this is given as:

$$F1 - Score = 2 \times \frac{Precsion \times Recall}{Precsion + Recall}$$

### 5.5 False Positive Rate (FPR)

The False Positive Rate (FPR) quantifies the ratio of benign cases misclassified as intrusions, reflecting the model's propensity to produce false alarms. The expression for this is given as:

$$FPR = \frac{FP}{FP + TN}$$

False Positives are benign instances misclassified as intrusion, while True Negatives are benign instances that are correctly identified. Reducing false positive rates is crucial for operational efficiency in intrusion detection systems, as excessive false positives can overwhelm security analysts and erode trust in the system.

## 6. Emerging Trends and Challenges

Machine learning (ML) has undergone significant advancements in recent years, changing the landscape of intrusion detection systems. The development of neural network models, like Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Generative Adversarial Networks (GAN), has proven instrumental in this transformation. These models can identify sophisticated attack patterns in high-dimensional network traffic with minimal feature extraction. There is growing interest in hybrid and semi-supervised training methods, which enhance detection accuracy and adaptability, particularly in environments where data is scarce or zero-day attacks are common. Additionally, there is a new focus on privacy-preserving strategies, such as Federated Learning, which enables intrusion detection collaboration among decentralized systems without compromising confidentiality.

Moreover, using explainable AI, such as SHAP analysis, can contribute to model transparency by highlighting the important features that affect predictions. A significant trend is the development of IoT-specific IDS designed for easy-to-interpret and protocol-independent intrusion detection in memory, power, and processing-constrained environments. Additionally, the use of sequential attention models enables the discovery of context- and time-dependent attacks. However, several challenges remain, including the high false positive rate, which continues to cause alert fatigue and erode the system's trust, thereby making IDS less effective. The lack of high-quality and properly labelled data hinders the use of supervised learning, and the computational requirements of many machine-learning methods are incompatible with the real-time nature of the problem. The black-box nature of deep learning methods does not satisfy the requirements of a security system in terms of interpretability and practicality.

Furthermore, the system needs to continuously adapt to new types of intrusions that static models cannot address due to their architecture. The resource constraints of edge and IoT devices limit the deployment of state-of-the-art machine-learning-based IDS systems. Finally, the reliance on obsolete and non-representative benchmark datasets limits the applicability and generalizability of these datasets in practice. Addressing these challenges is vital for developing an intrusion detection system that is robust, reliable, and practical.

## 7. Conclusion

The escalation in both complexity and the rate of cyber threat incidents requires an IDS that is intelligent and can dynamically adapt beyond traditional signature-based systems. ML enables the enhancement of IDS detection capabilities by automating threat detection procedures, recognizing patterns, and identifying zero-day attacks seamlessly.

This study comprehensively discusses different machine learning (ML) techniques, including supervised, unsupervised, semi-supervised, hybrid, and deep learning, as well as their attributes, constraints, and suitability for various implementations. The application of ML has enhanced detection accuracy, reduced false positives, and the ability to discover unseen and zero-day attacks.

Emerging innovations, such as federated learning, explainable AI, and deep learning architectures, are evolving the IDS paradigm, particularly in line with the shift to cloud, IoT, and edge computing models. Nevertheless, there are significant challenges, such as high false alarm rates, inadequate model interpretability, a lack of high-quality labelled data, and the tight resource restrictions of IoT and real-time systems. To address these challenges in the future, new research should focus on creating scalable, explainable, and dynamic models that utilize comprehensive, up-to-date datasets that reflect the latest threats. In conclusion, only a balanced application of intelligent detection methods and high-quality performance assessment metrics will successfully secure modern network infrastructures from emerging cyber threats.

**References**


1. AFRIDI, S. 2024. Machine Learning Innovations in Intrusion Detection Systems (IDS): Emphasizing Ensemble Learning for Enhanced Security.
2. AHMED, U., NAZIR, M., SARWAR, A., ALI, T., AGGOUNE, E.-H. M., SHAHZAD, T. & KHAN, M. A. 2025. Signature-based intrusion detection using machine learning and deep learning approaches empowered with fuzzy clustering. *Scientific Reports,* 15**,** 1726.
3. AKSHAYA, R. & SARAVANAN, C. A Novel Approach for Building Cyber Crime Prediction and Analysis Model using Random Forest. 2024 8th International Conference on Computational System and Information Technology for Sustainable Solutions (CSITSS), 2024. IEEE, 1-6.
4. AWAJAN, A. 2023. A novel deep learning-based intrusion detection system for IOT networks. *Computers,* 12**,** 34.
5. DASH, N., CHAKRAVARTY, S., RATH, A. K., GIRI, N. C., ABORAS, K. M. & GOWTHAM, N. 2025. An optimized LSTM-based deep learning model for anomaly network intrusion detection. *Scientific Reports,* 15**,** 1554.
6. DISHA, R. A. & WAHEED, S. A Comparative study of machine learning models for Network Intrusion Detection System using UNSW-NB 15 dataset. 2021 International Conference on Electronics, Communications and Information Technology (ICECIT), 2021. IEEE, 1-5.
7. DONG, H. & KOTENKO, I. 2025. Cybersecurity in the AI era: analyzing the impact of machine learning on intrusion detection. *Knowledge and Information Systems***,** 1-52.
8. ESHAK MAGDY, M., M MATTER, A., HUSSIN, S., HASSAN, D. & ELSAID, S. 2023. A Comparative study of intrusion detection systems applied to NSL-KDD Dataset. *The Egyptian International Journal of Engineering Sciences and Technology,* 43**,** 88-98.
9. FENG, J. Improved Machine Learning-based System for Intrusion Detection. 2024 2nd International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2024), 2024. Atlantis Press, 130-136.
10. KAVITHA, S. & UMA MAHESWARI, N. 2021. Network anomaly detection for NSL-KDD dataset using deep learning. *Information Technology in Industry,* 9**,** 821-827.
11. KHAN, M. & GHAFOOR, L. 2024. Adversarial machine learning in the context of network security: Challenges and solutions. *Journal of Computational Intelligence and Robotics,* 4**,** 51-63.
12. LIN, Q., LIU, Z., YANG, Y., WONG, K.-C., LU, Y. & LI, J. 2024. Multi-objective evolutionary neural architecture search for network intrusion detection. *Swarm and Evolutionary Computation,* 91**,** 101702.
13. MASEER, Z. K., YUSOF, R., BAHAMAN, N., MOSTAFA, S. A. & FOOZY, C. F. M. 2021. Benchmarking of machine learning for anomaly based intrusion detection systems in the CICIDS2017 dataset. *IEEE access,* 9**,** 22351-22370.
14. MOHAMMED, K. 2025. Enhancing Cybersecurity Through Artificial Intelligence: A Novel Approach to Intrusion Detection. *International Journal of Advanced Computer Science and Applications,* 16**,** 577 - 586.
15. MOUSTAFA, N. 2021. A new distributed architecture for evaluating AI-based security systems at the edge: Network TON_IoT datasets. *Sustainable Cities and Society,* 72**,** 102994.
16. OGUNBADEJO, M. D., OLUWATOBI A AYILARA-ADEWALE & MOGHADDAM, A. 2025a. A State-of-the-Art Review of Ransomware Attacks on Internet of Things: Trends and Mitigation Strategies. *Journal of Information Engineering and Applications,* 15**,** 1-18.
17. OGUNBADEJO, M. D., OLUWATOBI A. AYILARA-ADEWALE & ALADE, O. E. 2025b. Overview of Zero Trust Architecture Trend and Advancement in Information Security. *Journal of Information Engineering and Applications,* 15**,** 21-30.
18. RAI, H. M., PAL, A., ERGASH O'G'LI, R. A., UGLI, B. A. K. & SHOKIROVICH, Y. S. 2025. Advanced AI-Powered Intrusion Detection Systems in Cybersecurity Protocols for Network Protection. *Procedia Computer Science,* 259**,** 140-149.
19. SAHLI, Y. 2022. A comparison of the NSL-KDD dataset and its predecessor the KDD Cup'99 dataset. *International Journal of Scientific Research and Management (IJSRM),* 10**,** 832-839.
20. SHARIF, M. H. U. & MOHAMMED, M. A. 2022. A literature review of financial losses statistics for cyber security and future trend. *World Journal of Advanced Research and Reviews,* 15**,** 138-156.

21. TAYE, M. M. 2023. Understanding of machine learning with deep learning: architectures, workflow, applications and future directions. *Computers,* 12**,** 91.
22. THAKKAR, A. & LOHIYA, R. 2022. A survey on intrusion detection system: feature selection, model, performance measures, application perspective, challenges, and future research directions. *Artificial Intelligence Review,* 55**,** 453-563.
23. TRIPATHY, S. S. & BEHERA, B. 2024. A Review of Various Datasets for Machine Learning Algorithm-Based Intrusion Detection System: Advances and Challenges. *IJISAE,* 12**,** 3833-3857.
24. WANG, Q., XIE, M., WU, Z. & YANG, D. Network Intrusion Detection and Dynamic Defense Method Based on Unsupervised Machine Learning. 2023 International Conference on Computer Simulation and Modeling, Information Security (CSMIS), 2023a. IEEE, 75-80.
25. WANG, Y., ZHENG, W., LIU, Z., WANG, J., SHI, H., GU, M. & DI, Y. 2023b. A federated network intrusion detection system with multi-branch network and vertical blocking aggregation. *Electronics,* 12**,** 4049.
26. YADAV, R., PATHAK, P. & SARASWAT, S. 2020. Comparative study of datasets used in cyber security intrusion detection. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology,* 6**,** 302-312.
27. ZUKAIB, U., CUI, X., ZHENG, C., HASSAN, M. & SHEN, Z. 2024. Meta-IDS: meta-learning based smart intrusion detection system for internet of medical things (IoMT) network. *IEEE Internet of Things Journal*.