MACHINE LEARNING FOR INTRUSION DETECTION: TRENDS, CHALLENGES AND FUTURE DIRECTIONS

*Resmi Krishnan V * [1], S. Mythili [2]*

ABSTRACT

Networks and information systems must be protected from cyberattacks by intrusion detection systems. Traditional rule- based intrusion detection systems are becoming less and less effective against complex and dynamic threats. Consequently, machine learning (ML) approaches are being used to create more intelligent, flexible and instantaneous threat detection systems. The Machine learning approaches utilised in IDS are supervised, unsupervised, ensemble, deep learning and federated learning models which is thoroughly reviewed in this work. Their benefits, drawbacks and practical uses are highlighted. This study also covers performance evaluation measures, datasets and new developments including edge computing solutions, privacy- preserving intrusion detection systems and Explainable AI (XAI). The study offers possible remedies while shedding light on problems related to feature selection, class imbalance, scalability and false positives. This research aims to provide researchers and practitioners with a better understanding of recent advancements in IDS models.

Keywords: Intrusion detection system, Ensemble learning, Machine learning, Deep learning, Federated learning, Explainable AI.

I. INTRODUCTION

Since cyber attacks are becoming more frequent, vast and complicated protecting computer networks and information systems is a challenging issue in this digital era. Intrusion detection systems (IDS) are essential component of modern cyber security framework because they can only detect and prevent unlawful activities, policy violations, and potential network attacks. Traditional intrusion detection systems (IDS), which mostly rely on signature-based anomaly detection find it challenging and nearly failed to keep up with evolving threats like advanced persistent attacks (APT) and zero-day exploits. Due to this changing cyber threat landscape, there has been a surge in the use of machine learning (ML) algorithms that provide intelligent and adaptable capabilities to enhance the effectiveness of IDS.

Machine learning based intrusion detection systems utilize capacity to automatically recognize patterns in data and spot variations that can point to a malicious activity. These algorithms fall into three categories mainly ensemble, supervised, and unsupervised. Labelled datasets are used to train models in supervised approaches however lack their ability to identify invisible threats. On the other hand, unsupervised methods are appropriate for unidentified threats since they identify irregularities without requiring prior knowledge of attack patterns. Ensemble techniques overcome the limitations of individual models by combining several algorithms to increase detection accuracy. Despite these benefits Machine Learning - based intrusion detection systems encounter difficulties with feature selection, high false positive rates, scalability and real- time performance.

General Educational Department[1]
GGVHSS School, Nemmara, Palakad[1]
resmikishor2010@gmail.com[1]
Department of Computer Science[2],
Karpagam Academy of Higher Education, Coimbatore, India[2]
smythili78@gmail.com[2]

* Corresponding Author

Table 1: Machine Learning Algorithms In Ids

YEAR	NUMBER OF STUDIES/PAPERS USING ML TECHNIQUES	KEY MACHINE LEARNING METHODS HIGHLIGHTED
2019	20+	Traditional ML algorithms (e.g., SVM, Decision Trees, k-NN), initial hybrid models
2020	30+	Deep learning models like DBNs, CNNs, and autoencoders; focus on anomaly detection.
2021	40+	Federated learning, IoT-specific IDS, emphasis on feature engineering.
2022	50+	GANs, ensemble methods, real-time detection, cloud-based ML approaches.
2023	60+	Explainable AI (XAI), hybrid deep learning models, privacy-preserving IDS.
2024	70+	Advanced multi-layer neural networks, 5G and edge-computing IDS solutions.

With an emphasis on significant developments, constraints and unresolved research issues, this work attempts to present a thorough analysis of machine learning methods used in Intrusion Detection System. This study focusses on assessment metrics, different machine learning models such as federated learning, hybrid models, and deep learning and assess how well different approaches perform on benchmark datasets. This study aims to provide academicians and practitioners with useful insights by synthesizing data from previous studies, directing futureadvancements towards more reliable, flexible, and effective Intrusion Detection System solutions.

II. RELATED STUDIES

Due to increasing complexity of cyber-attacks, there has been a tremendous increase in the use of machine learning (ML) techniques in Intrusion Detection Systems (IDS) in recent years. Numerous studies have investigated various machine learning techniques such as supervised, unsupervised and ensemble methods to enhance IDS performance. Due to their superior ability to handle complex network data and invisible malicious attacks Deep learning (DL) models like auto encoders and convolutional neural networks (CNN) are becoming more popular. Initially, researchers concentrated on traditional Machine Language classifiers like decision trees, support vector machines, and k-nearest neighbors.

The use of machine learning in Intrusion Detection System (IDS) to solve security and privacy related issues on the Internet of Things is covered in the research [1]. Concept drift, high dimensional data and computational complexity are the major drawbacks with traditional intrusion detection models. The authors suggested that machine learning approaches is the only solution for such problems. The popular datasets examined by them for IDS development are Kyoto, NSL-KDD, and KDD99 highlighting the necessity of striking a balance between detection accuracy and computing efficiency in order to adjust to changing IoT contexts.

[2] compared different machine learning classifiers for intrusion detection systems (IDS). Moreover, this paper provides valuable information on how well the classifiers perform real time in terms of recall, accuracy, precision and training time. [3] describe a method for using Deep Belief Networks (DBNs) for anomaly detection in the Internet of Things (IoT).

An intrusion detection system (IDS) that uses auto encoders for accurate anomaly detection is presented in the research [4]. The study [5] suggested a hybrid intrusion detection system model for IoT environments which combines machine learning with anomaly-based and signature-based detection.[6] studied a modified clustering approach for improving intrusion detection framework. [7] studied the effectiveness of many deep learning architectures, including recurrent neural networks (RNNs) and convolutional neural networks (CNNs), at detecting intricate intrusion patterns in network traffic data.

The study by [8] focused on federated learning approaches for intrusion detection systems (IDS) in distributed network environments. The study by [9] examines ensemble learning techniques for network intrusion detection and proved that ensemble methods can effectively reduce false positives and improve detection rates. The study by [10] aimed at developing intrusion detection systems based on adaptive learning strategies such as ensemble learning for improving network security.

[11] studied about real time intrusion detection system using deep learning techniques. Their study highlighted that real-time intrusion detection systems can be used for quickly identifying intrusions and deep neural networks can process massive data streams quickly with accuracy.

The study conducted by [12] provided organizations looking for real time security solutions for larger networks in cloud computing environment. The goal of [13] study is to assess classifiers like Decision Trees, Support Vector Machines and Neural Networks on various datasets to identify network hazards.

[14] study presents a unique intrusion detection system that uses auto encoder networks to detect unusual activity within network data. A comprehensive survey of supervised learning methods in intrusion detection system is given by [15] describe that classifier like Support Vector Machine, Random Forests and Neural Networks are employed to identify unknown attack patterns. [16] discussed the development of intrusion detection system designed specifically for edge computing environments.

Standardized datasets are essential for enhancing the effectiveness of intrusion detection systems according to study by [17]. For smart grids [18] presented a hybrid intrusion detection technique that increases detection precision. The hybrid model uses classification with Long Short-Term Memory (LSTM) networks and feature extraction with XGBoost and Convolutional Neural Networks (CNNs). Generative Adversarial Networks (GANs), which create possible attack scenarios and model data distributions was used by [19] to present an anomaly-based intrusion detection system.

The research by [20] examines feature reduction techniques for real-time intrusion detection systems, including methods like Principal Component Analysis (PCA) and auto encoders to enhance performance by lowering computational demands while maintaining accuracy. [21] study underscore the versatility of deep learning and ensemble methods across different network environments, especially in reducing false positives.

[22] study investigates the deployment of AI-based intrusion detection systems (IDS) specifically for mobile networks. The research by [23] and colleagues addresses the challenges intrusion detection systems face, especially with the growing complexity of cyber threats. [24] work reviews a deep learning- based IDS developed for the Industrial Internet of Things (IIoT).

[25] paper discusses the application of machine learning techniques in IDS for 5G networks. The importance of explainable artificial intelligence (XAI) in intrusion detection is underscored by study [26] which assists security analysts in understanding and evaluating decisions made by complex Machine Learning-based IDS models. The research by [27] discusses the array of data attributes and types of attacks that can influence IDS efficiency, focusing on tailored performance metrics for evaluating IDS across different datasets. In [28] research delve into cutting-edge developments in machine learning (ML)-based intrusion detection, including self-learning models, federated learning, and integration with 5G and IoT networks. The study by [29] sheds light on Deep reinforcement learning-based IDS, optimizing the model's ability to adapt to changing network behaviors and autonomously adjust to new types of intrusions. In [30] research explores federated learning as a privacy-preserving approach in intrusion detection systems (IDS), allowing multiple entities to collaborate on improving model performance without sharing sensitive data.

III METHODOLOGY

Using a comparison technique, this review paper examines different machine learning (ML) algorithms utilized in intrusion detection systems (IDS) based on Machine Learning approaches.It examines important machine learning techniques including supervised, unsupervised and hybrid approaches evaluating how well they identify anomalies and harmful activities. This evaluation provides feasibility study of each model for real time detection, idea drift, flexibility and scalability within network, recall and computing efficiency.

A. DATA COLLECTION AND FEATURE EXTRACTION

Finding peer reviewed publications and conference papers about machine learning techniques for intrusion detection systems (IDS) is a part of the data collection process for this review paper. The selection process uses relevance to machine learning IDS methodologies, publication dates from 2019 to 2024 and impact measures to filter studies. This procedure aims to ensure a comprehensive and current representation of developments in machine learning for IDS technology.

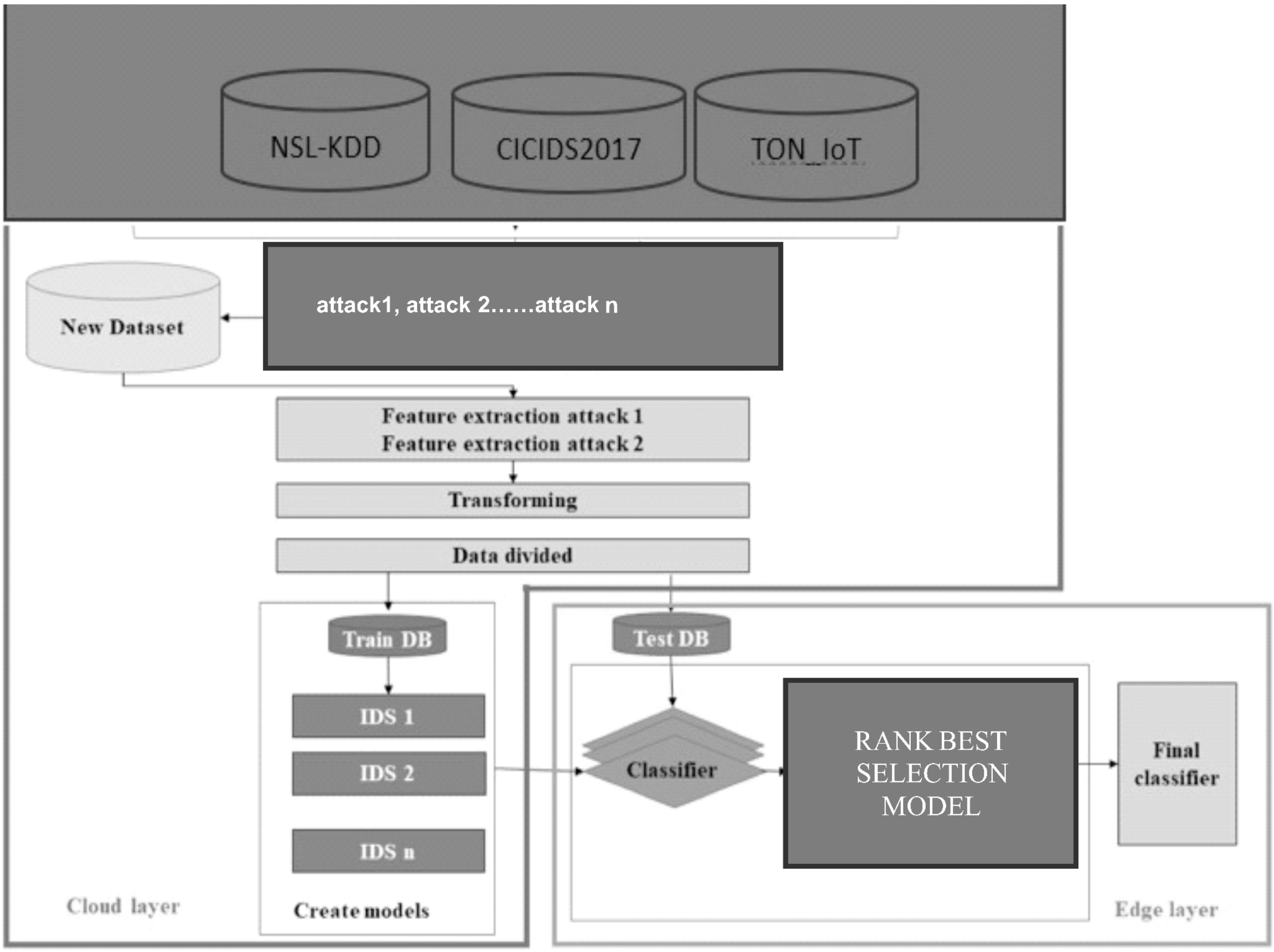

Figure 1 : Ids Architecture With Ml Integration

To ensure robust IDS model evaluation, researchers utilize various benchmark datasets.

Table 2: Overview of Datasets Used in the Study

Dataset	Year	Features	Attack Types	Application Domain
NSL-KDD	2009	41	DoS, U2R, R2L, Probe	Network IDS
CICIDS2017	2017	80+	DDoS, Brute Force	Real-time traffic

Feature selection methods play a crucial role in enhancing the efficiency and effectiveness of IDS models:

1. Principal Component Analysis (PCA) - Principal Component Analysis (PCA) converts high-dimensional data into a lower- dimensional space with most of the variation retained. The model performance is enhanced further by eliminating redundant features. PCA reduces dimensionality by transforming the dataset into a set of linearly uncorrelated principal components:

$$C = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})^T \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

where x_i represents the feature vectors, and $\bar{x}$ is the mean feature vector.

The principal components are obtained by solving:

$$c_v = \lambda_v \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

where λ are the eigen values and v are the corresponding eigenvectors.

2. Correlation-based feature selection - The most relevant features are retained by applying correlation-based feature selection, which identifies and removes highly correlated features. By reducing multicollinearity correlation-based feature selection enhances IDS model accuracy and interpretability. This method measures the correlation between input features and the target variable using the Pearson correlation coefficient:

$$r_{xy} = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2\sum(y_i-\bar{y})^2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

where x_i and y_i are individual observations of features and the target variable.

3. Mutual Information- Mutual information measures how much the target variable depends on the input features. Features with a higher mutual information score are selected for model training since they give a larger contribution to intrusion detection. Identification of non-linear relationships among variables is facilitated by this method.

$$I(X;Y) = \sum_{x\in X}\sum_{y\in Y} p(x,y)\log\frac{p(x,y)}{p(x)p(y)} \quad \ldots\ldots (4)$$

where $p(x,y)$ is the joint probability distribution of features x and target y

4. Recursive Feature Elimination (RFE)- In the dataset, all but the most important features remain due to the use of Recursive Feature Elimination (RFE), which has the effect of eliminating the less important features incrementally by examining model

performance. This method takes particular advantage when used with supervised learning models.

Using these feature selection approaches, IDS models can improve detection accuracy, minimize computational complexity and avoid over fitting. Constructing IDS models involves selecting relevant features and dataset quality. Dimensionality reduction and enhanced classification performance are often achieved using Principal Component Analysis (PCA), Mutual Information and Recursive Feature Elimination (RFE). As typical datasets for IDS testing, NSL-KDD and CICIDS2017 offer diverse attack scenarios for training and testing Machine Learning models.

B. CATEGORIZATION OF MACHINE LEARNING TECHNIQUES

Supervised learning, unsupervised learning and hybrid methods are the three main categories into which the reviewed literature divides machine learning techniques for intrusion detection systems. Due to the rapid increase in cyber threats, there has been a shift toward more complicated models like deep learning and federated learning. The quantity of papers, the variety of algorithms and application domains including cloud environments and Internet of Things are the main subjects of analysis. Studies compare the efficiency of each method by using performance criteria such as accuracy, precision, recall, F1-score and latency.

SUPERVISED LEARNING

Using labelled datasets, supervised learning approaches train intrusion detection systems (IDS) models to identify malicious or benign network traffic. Typical algorithms for supervised learning are:

- ❖ Support Vector Machines (SVM) -are computationally costly but effective with high-dimensional data.
- ❖ Decision trees (DT)- are quick and easy to understand, yet they might overfit.
- ❖ Random Forest (RF)-An ensemble of decision trees called Random Forest (RF) lowers variance and increases accuracy.
- ❖ Neural Networks (NN)- Deep learning models that can identify intricate attack patterns, but they demand a lot of processing power.

UNSUPERVISED LEARNING

Network traffic anomalies can be found using unsupervised learning techniques without the need for labelled data. Typical methods include of:

❖ Clustering (K-Means and DBSCAN)-Network traffic is grouped into clusters according to similarities.

❖ Autoencoders: Deep learning models that identify anomalies by identifying departures from typical traffic patterns.

❖ Isolation forests-An efficient tree-based technique for identifying outliers in high-dimensional datasets

ENSEMBLE LEARNING

Efforts directed at ensemble learning aim at advancing model robustness and accuracy by combing a range of models. This involves a number of techniques.

❖ Bagging (Random Forests)-Ensemble learning algorithms proceed to build a collection of learners and then employ them for training. Such trees can be built on different bootstrap versions of the training data set.

❖ Boosting (such as AdaBoost and XGBoost)- In this case, weaker sub-learners are integrated in order to obtain a stronger learner. It is based on the idea that we build learners one at a time, and each new learner tries to fix the mistakes that the right model did for your predictions.

❖ Stacking – This is the super advanced ensemble method where several machine learning models are combined and the final predictions are made by a meta-learner.

C. OPTIMIZATION TECHNIQUES FOR IDS MODELS

IDS Performance enhancements through optimization strategies goes hand in hand using the following techniques-

1. Hyperparameter Tuning-As part of model optimisation, model parameters can be fine-tuned such that performance significantly improve. Methods commonly include:

❖ Grid Search: Over a specified value of the hyper parameters.

❖ Random Search: Best hyper parametric set is achieved by choosing random.

❖ Bayesian Optimisation: Maximizing the rate of performance using some probability models.

2. Cross-Validation

Cross-validation guarantee dependable model performance and avoids overfitting. Principal techniques include:

❖ K-Fold Cross-Validation: Disintegrates the data into K parts and trains the model on different configurations.

❖ Stratified Cross-Validation: It guarantees that there is no imbalance in the class distribution in the training and validation sets.

3. Dropout Regularization

Dropout is applied to deep learning structures to ensure that there is no over fitting by turning off a certain random percentage of all neurons during their operation. It is also one way of making generalized and robust models.

D. EVALUATION OF MODEL EFFECTIVENESS

To assess the effectiveness of the model in various machine learning based intrusion detection systems key performance indicators are measured and analyzed. Accuracy measures the ability to correctly identify malicious and benign traffic. Recall assesses the model's sensitivity to actual intrusions which is crucial for reducing false negatives. Precision concentrates on the true positive rate among identified intrusions. For datasets with class imbalances the F1-Scoreoffers a balanced perspective of model performance by combining precision and recall. Latency assesses the capacity for real-time processing which is crucial for systems that need to react quickly to threats.

TABLE II. PERFORMANCE METRICS OF ML MODELS IN IDS

Model Type	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Latency (ms)
Supervised Learning	85	83	80	81	20
Unsupervised Learning	78	75	70	72	15
Hybrid/Ensembe Models	90	88	85	86	25

IV. CHALLENGES IN ML-BASED IDS

Optimising model performance while maintaining scalable and real-time threat detection is the first challenge in Machine Learning based intrusion detection systems. Effective model training is hampered by the challenges associated with feature selection in high-dimensional datasets. Attack samples that are under-represented due to class imbalance may result in low detection rates for uncommon but serious threats. Another serious challenge faced by machine learning based intrusion detection system is false positives and such incorrect classifications might result in pointless warnings that hinder security efforts. Scalability is also a major concern since large- scale network settings require adaptable and resource-efficient machine learning models to manage growing traffic volumes without sacrificing performance.

V CONCLUSION AND FUTURE RECOMMENDATIONS

This paper emphasizes the significant advancements and trends in machine learning-based intrusion detection systems underscoring their growing relevance in addressing complex and evolving cyber security threats. The investigation demonstrates the effectiveness of various machine learning techniques ranging from traditional supervised methods to more sophisticated hybrid and ensemble approaches. These techniques offer enhanced accuracy, flexibility and scalability for real-time intrusion detection systems. It is discovered that cloud-based and edge computing solutions are crucial for managing environments with high traffic and limited resources, which makes IDS more flexible to meet the needs of contemporary networks. Additionally, it is clear how crucial it is to pick and assess a variety of datasets in order to provide robust model evaluation and guarantee generalizability across real-world applications. Overall, advancements in computational infrastructure and ongoing ML algorithm development promise to substantially bolster IDS capabilities, making them indispensable in today's cyber security landscape. According to the analysis deep learning models such as recurrent and convolutional neural networks are becoming more and more well-liked in IDS because to their accuracy and versatility. However, in situations with limited resources such as in the case of Internet of Things their high computing demand presents deployment issues. For lightweight Intrusion Detection

System applications conventional machine learning methods like decision trees and support vector machines are still applicable. The rise of hybrid and ensemble methods which merge the strengths of multiple models to enhance detection precision while addressing concept drift and evolving network threats represents a significant trend. As machine learning based Intrusion Detection System models became more advanced, they have overcome the drawbacks of traditional methods by enhancing processing speed, accuracy and adaptability. However, challenges remain in the areas of datasets availability and standardization particularly for emerging fields such as IoT and 5G. Future directions include integration of federated learning for privacy-preserving Intrusion Detection System and exploration of explainable Artificial Intelligence to enhance system transparency and trust.

1. Explainable Artificial Intelligence (XAI) For Intrusion Detection Systems-Explainable Artificial Intelligence (XAI), which has emerged in the recent trends, is essentially responsible for designing ML-based Intrusion Detection Systems that are human dignifying in nature, as it is the case the norm. Such

technologies are useful as will be best explained further in this part are. With this, it can be quite hard explaining to a security analyst how a particular attack has been caught as in many traditional ML or deep learning models operate as black boxes and are therefore near impossible to trace. There are numerous methods of XAI applications obtainable today including; SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) which can be applied to explain the reasons for a machine's decision. As a result, this explains why greater trust, responsibility, and the ability to debug have security systems in relation to malicious threats. In Progressing IDS technologies such as these vulnerable technologies it will be obligatory to introduce XAI for the purpose of improved clarity in decision making, compliance as well as the trust of the end user.

2. Privacy-Preserving Federated Learning: Federated learning allows for training an intrusion detection system in a coordinated manner without putting data confidentiality at risk. It would be our recommendation to implement optimized federated learning in the design and structure of IDS in upcoming systems more so in protected areas.

3. Edge and Cloud-Based Intrusion Detection Systems Put up an IDS at the edge destination will reduce detection latency and make the threat retaliation in real-time a reality. Management of cloud-based IDS solutions will require the employment of real time machine learning solutions.

4. Adaptive Learning with Reinforcement Learning Reinforcement learning is incorporated in the IDS systems, it would enhance self-optimizing toward the changing cyber space. Such predictive systems can automatically upgrade detection mechanisms upon scanning all new attacks.

5. Development of Comprehensive Benchmark Datasets Tangible inputs as benchmarks must be quantified for future intrusion detection system research. More specifically such benchmarks should also be supplemented by standardization of data for the training of models in diverse networks in future attempts of evaluations.

The ongoing advancements in hardware such as edge computing also present promising avenues for deploying these sophisticated IDS models in diverse and real- world settings. The advancementof federated learning, particularly for privacy-sensitive contexts like IoT and 5G where centralized data storage is not feasible represents a significant direction for the development of Machine language-based Intrusion Detection Systems (IDS). It is important to integrate adaptive learning features that enable the IDS to respond to new attack vectors in real time, especially through reinforcement learning. Researching lightweight, efficient models suitable for edge computing is very essential to guarantee performance in limited-resource settings. Lastly, the creation of definitive benchmarks and varied datasets is imperative for consistent and dependable evaluations and thus boosting the effectiveness of IDS under various network scenarios.

REFERENCES

[1] Ghani, A. A. A., Abdullah, A., & Hakim, F. (2021). An IDS for IoT using machine learning. MDPI.

[2] Khan, F. A., & Gumaei, A. (2019). Machine learning classifiers for intrusion detection. Springer.

[3] Thamilarasu, G., & Chawla, S. (2020). Anomaly detection for IoT using DBNs. IEEE Xplore.

[4] Sandeep, G., et al. (2020). Sparse autoencoders for IDS. IEEE Xplore.

[5] da Costa, K. A. P., et al. (2021). Hybrid IDS models in IoT. Cyber security .Journal.

[6] Yang, X., et al. (2019). Modified clustering for IDS. arXiv.

[7] Gupta, M., et al. (2022). Advances in deep learning-based IDS. IEEE Xplore.

[8] Ali, R., & Singh, S. (2023). Federated learning approaches for IDS. Springer.

[9] Sharma, A., et al. (2024). Ensemble learning for network intrusion detection. IEEE Transactions on Cybernetics.

[10] Oprea, M., et al. (2021). A survey on evolving IDS and concept drift handling. Frontiers in Computer Science.

[11] Patel, N., & Joshi, A. (2022). Real-time IDS with deep learning. Journal of Network Security.

[12] Pande, A., & Chauhan, R. (2023). Scalable IDS using cloud-based ML. MDPI.

[13] Li, Y., et al. (2019). A comparative study of ML algorithms in IDS. IEEE Access.

[14] Wang, J., & Zhang, Q. (2020). Security-aware IDS using autoencoders. Symmetry.

[15] Malik, S., et al. (2021). Survey on supervised IDS methods. Springer.

[16] Acharya, B., et al. (2022). Resource-efficient IDS in edge computing. IEEE Internet of Things Journal.

[17] Han, T., & Liu, F. (2020). Benchmarking datasets for IDS. Journal of Big Data.

[18] Park, J., & Kim, H. (2023). Hybrid intrusion detection for smart grids.Energy Informatics.

[19] Bansal, R., et al. (2022). Anomaly-based IDS using GANs. IEEE Access

[20] Xiao, Y., et al. (2021). IDS with feature reduction techniques. Journal of Cybersecurity.

[21] Singh, P., & Kumar, S. (2023). Comparative analysis of anomaly detectors. Computer Networks.

[22] Rana, A., et al. (2024). AI-based IDS for mobile networks. Springer.

[23] Farooq, A., et al. (2019). Review on IDS challenges and solutions. IEEE Xplore.

[24] Desai, K., & Vora, M. (2020). IDS for industrial IoT using deep learning. MDPI.

[25] Haque, T., & Rahman, M. (2022). Real-time intrusion detection in 5G. Journal of Network Security.

[26] Roy, D., et al. (2023). Explainable AI in IDS. IEEE Transactions on Cybernetics.

[27] Ahmed, F., & Hussain, Z. (2021). Dataset-specific IDS performance metrics. Journal of Computer Science.

[28] Zhao, J., et al. (2024). Future trends in ML-based IDS. IEEE Access.

[29] Neupane, K., & Shrestha, P. (2019). IDS with deep reinforcement learning. Journal of Cyber security.

[30] Joshi, A., & Patel, R. (2023). Federated IDS with privacy- preserving techniques. IEEE Access