Multi-Class Attack Detection in Wireless Sensor Networks Using Ensemble Learning Techniques

Bipin Bihari Jayasingh
IT Department
CVR College of Engineering
Hyderabad -501510, India.
Email: bipinbjayasingh@cvr.ac.in

M Yogitha
IT Department
CVR College of Engineering
Hyderabad – 501510, India.
Email: mittayogitha123@gmail.com

***Abstract* — Wireless sensor networks in various types of infrastructures are increasing rapidly, especially as part of Industry 4.0, which necessitates real-time monitoring and control in a fully automated manner in industrial, agricultural, healthcare and smart grid environments. However, due to the nature of wireless communication and sensor network device constraints, they are targeted by cybersecurity attacks such as Blackhole, Grayhole, Flooding and time-division multiple access (TDMA) shipping attacks that disrupt routing, consumption of available channel bandwidth and in worst-case scenarios. This paper describes a comparative performance evaluation of an ensemble-based intrusion detection system on the WSN-DS benchmark dataset. The WSN-DS dataset consists of 374,661 labelled records of network traffic at the network layer and includes both normal traffic as well as four attack types. The dataset contains seven classical machine learning algorithms including Decision Tree, Random Forest, Support Vector Machine, Naive Bayes, K-Nearest Neighbors, Logistic Regression and XGBoost a feature set of 19 original features resulting in 14 of the most discriminative features combined with a corresponding data transformation technique (z-score normalisation) and the application of a five-fold stratified cross-validation evaluation approach that yields reliable and repeatable accuracy estimates suitable to account for the exceedingly high degree of class imbalances associated with this dataset. The results demonstrate that the overall accuracy of the ensemble approach is 99.5%.**

Keywords*— *Wireless Sensor Networks, Intrusion Detection System, Ensemble Learning, Majority Voting, Random Forest, XGBoost, WSN-DS, Cyberattack Classification, Industry 4.0, Feature Selection.

I. INTRODUCTION

The increase of wireless sensor networks in industrial, health care, agriculture and military sectors has been one of the most significant technology trends in the last 20 years [7]. Numerous sensor nodes make up wireless sensor networks. spread out over an area, collecting data from the physical environments that are transmitted back to a central processing unit using wireless communication [8].

In the scope of Industry 4.0 Wireless sensor network, WSNs function as the central nervous system for smart factories and connected infrastructure by providing real-time visibility into operational conditions and enabling automated responses [16]. While wireless sensor networks have benefits by supporting numerous applications, they also have vulnerabilities that arise from being deployed in open-air wireless environments [3].

This type of deployment leaves the sensor nodes vulnerable to malicious interference [5]. Wireless Sensor nodes have extremely little memory, computing power, and battery life capacity [6]. As a result, security mechanisms common to enterprise networks are not practical for deployment in networks of wireless sensors [9]. Taking advantage of the security holes mechanisms allows an attacker to launch targeted attacks on the wireless sensor networks, causing disruption in packet routing, flooding the communication channel, or injecting false scheduling information [4].

The effects of these attacks may create serious problems [13]. Within an industrial environment, if corrupted sensor information is automatically responded to, it may lead to equipment not functioning properly or critical anomalies going undetected [16]. Similarly, if a wireless sensor network is compromised and contains medical data, it could endanger the lives of patients in a medical monitoring system [14]. Therefore, there is a need for research on the creation of reliable and lightweight intrusion detection systems (IDS) that satisfy the requirements of wireless sensor networks, while balancing the trade-offs existing within each strategy and providing guidelines on how each can be used in practice in the resource-constrained sensor environment [2],[11],[15].

II. RELATED WORK

As the number of Wireless Sensor Networks facing potential cyber threats continues to rise, there has been significant research into utilizing machine learning techniques for intrusion detection within these networks [1]. This work introduces a situationally aware predictive model designed to identify and prevent cyber intrusions at the entry points of wireless sensor networks, particularly those deployed in the context of Industry 4.0. The data gathered from real-time sensor networks is analyzed using both deep learning and machine learning techniques to better identify unusual activities and potential malicious attacks within the WSN, aiming to strengthen their ability to detect such issues as they happen. The model's use of context-specific information and adaptive learning methods also enhances its capability to accurately identify a wide variety of harmful events and take steps to prevent potential threats from developing into actual problems. The following assessment shows that using predictive analytics along with smart models creates a safer environment for WSNs against continuously changing cyber threats.

Kurtkoti et al. [2] were one of the first researchers to conduct work in this area when they presented a framework for assessing machine learning techniques for identifying flooding, blackhole, and grayhole attacks in WSNs, utilizing the WSN-DS dataset for their analysis. Their researchers compared several classifiers across various train-test splits, reporting that the best-performing classifier across all three types of attack was AdaBoost; therefore, helping to establish a benchmark upon which to base future work in this field.

Sadeghi and Alzubaidi [3] provides an overview of machine learning methods that have been utilized in Network Intrusion Detection Systems (NIDS). Based on an analysis of the literature, the study finds a number of challenges associated with the design of IDS. In particular, the authors describe two major challenges related to the design of IDS: (I) the existence of imbalanced datasets, and (II) the need to have an explainable model of AI. The authors acknowledge that their review did not provide experimental evidence to support the findings they have reported.

To address the class imbalance issue seen in previous surveys, Talukder et al. [4] suggested a hybrid model for intrusion detection called the MLSTL-WSN model, which implemented the SMOTE-Tomek technique for class balancing between attacks and normal/legitimate traffic classes in the dataset. After applying this technique to create a balanced dataset, multiple classifiers (K-Nearest Neighbors, Random Forest, Support Vector Machine, and Decision Tree) were trained on that dataset which allowed for the identification of rare attacks that may otherwise go undetected in WSNs, thereby proving that preprocessing methods are just as important as the choice of classifier.

Arabiat et al [5] have created an intrusion detection system that uses machine learning techniques to identify and categorize denial of service (DoS) threats against heterogeneous wireless sensor networks (WSNs) utilizing the WSN-DS data set as input to the classifier algorithms being implemented such as AdaBoost, CN2 Rule Inducer, and Random Forest. The authors used preprocessing and normalisation methods before training the models. They provide empirical evidence to support their hypothesis regarding the effectiveness of preprocessing data before conducting classification tasks; the authors report that their use of AdaBoost achieved a 100% classification rate for multiple DoS attack types, including Flooding, Scheduling, Blackholes, and Grayhole. This supports the assertion that appropriate data preparation can provide reliable detection performance.

Support Vector Machine, Random Forest, Naive Bayes, and Convolutional Neural Networks are only a few of the machine learning techniques used in the overall anomaly-based intrusion detection method developed by Ugale et al. [6]. They used common metrics including accuracy, precision, recall, and F1 score to evaluate the model's efficacy, and they showed that the results highlight the value of using machine learning algorithms for detecting abnormal network activity across various threat scenarios.

In agreement with previous research showing that ensembles outperformed other models, Sharma et al. [7] developed a modelling framework based using machine learning methods to detect Wireless Sensor Network (WSN) Denial of Service threats through a data-driven approach to intrusion detection. This work employed two characteristics of mathematical modelling, preprocessing and feature evaluation to transform data into a form suitable for machine learning models and working within collaborative groups of similar sensors to achieve global, as opposed to individual, detection accuracy. Their findings confirmed that ensemble classifiers produced higher detection rates (increased true positives) and lower false-positive rates than individual classifiers, further demonstrating the advantages of using aggregated models in resource-limited environments.

Kumar and Patel [8] designed a lightweight intrusion detection system designed specifically for wireless sensor networks with little resources, understanding that

real-world WSN deployments have computational limitations. Using classifiers like Decision Trees and Support Vector Machines to identify malicious nodes, the authors demonstrated that using machine-learning models to detect abnormal traffic patterns could be achieved with minimal computational resources so that there exists a balance between detection accuracy and practical deployability.

S. Jadhav et al. [9] utilized machine learning techniques as a means to detect network intrusions and express the value of using classification algorithms for the recognition of malicious activities in network traffic. Both of these studies highlight the significant. The impact of feature selection and model optimization on detecting malicious activity more accurately and reducing false positives that occur.

A machine learning-based intrusion detection system [10] assessed machine learning methods for application in identifying intrusions in modern networks and emphasized the relevance and flexibility of models that were developed for detecting different types of attacks. Both of these research efforts concluded that advances in machine learning methods are an important tool for the improvement and reliability of intrusion detection systems, especially in a large-scale and dynamic network environment.

Building on hybridisation, as discussed previously, Nandhakumar [11] gives us a better version of the effect of feature selection and model optimization on improving the accuracy of harmful activity detection and lowering false positives adaptive and hybrid development algorithms, using machine learning techniques as part of its architecture. By utilizing a hybrid versus single algorithm approach, the work shows increased accuracy in detecting attacks when used in dynamic environments where attack patterns will continue to change over time. This further substantiates that static models will not be adequate for long-term deployments of WSN.

As technological improvements have been made in detecting accuracy, there has been an increase in transparency of model outputs as well as accountability of machine learning methods that provide the basis for systems that make decisions automatically. To deal with these deficiencies, Patil et al. [12] have developed a framework that provides interpretable classification results and an explanation for those results by using "Explainable Artificial Intelligence" (XAI) methods to support intrusion detection. This framework can support regulated and safety-critical applications, in addition to fulfilling the requirement of balancing predictive effectiveness with human-understandability of the logic used to support each AI-based output.

In addition to this, researchers have begun to apply deep learning techniques along with more sophisticated optimization methods to improve detection performance. For instance, Garcia and Blandon [13] an intrusion detection and prevention solution for denial-of-service attacks based on deep learning achieving good results according to IEEE Access benchmarks. Also, Pradeepa and Ponnusamy [14] improved WSN security by combining recurrent neural networks and optimization algorithms to provide better temporal pattern recognition of sequential attack behaviours based on sequential modelling. Finally, Turukmane and Devendiran [15] introduced M-MultiSVM, an SVM (Support Vector Machine) framework for multi-class classification that incorporates feature selection to decrease the number of features and increase classification accuracy
required. This work reinforced that intelligent feature engineering is an important tool for enhancing intrusion detection systems' (IDSs') functionality.

III. METHODOLOGY

Fig 3.1. Machine learning pipeline for intrusion detection using the WSN-DS dataset.

A. Dataset

The WSN-DS benchmark dataset is used in all experiments and was gathered from a simulated WSN environment that mimics the network conditions of Industry 4.0. The benchmark consists of 374,661 records, which are classified into one of five classes Normal, Blackhole, Grayhole, Flooding, or TDMA. The records comprise 19 network-level features, including counters for advertisement packet (ADV_S, ADV_R), join request (JOIN_S, JOIN_R), scheduling (SCH_S, SCH_R), data packet (DATA_S, DATA_R), residual energy, size of data transmitted and received, distance to the cluster head, and identifiers of the cluster head. This dataset is publicly available and has been used as a benchmark in the literature on WSN intrusion detection; it is therefore suitable for use in reproducible studies. The class distribution is heavily unbalanced: 93% of records are for normal traffic, whereas the four attack classes make up the remaining 7% (with Flooding being the least frequent). This is a critical issue, as a classifier that predicts every record as Normal would get an apparent accuracy of more than 93%, but it would provide no protection. To counter this, we evaluate results using the macro-averaged precision, recall and F1-score metrics, as well as accuracy, and we use stratified cross-validation to ensure that all folds have the same class distribution.

B. Preprocessing and Feature Selection

There are three stages of raw data processing. First, we filter out records with missing (or very outlying) sensor readings (more than five standard deviations from the mean value of the feature) (less than 0.1% of the data). Second, all continuous features are normalised to zero mean and unit variance by z-score normalisation, using statistics from the training fold and applied to the test fold, to avoid data leakage. Third, categorical features (cluster-head IDs and send_code) are converted to integers. We use Pearson correlation for feature selection. Features with absolute correlation below 0.05 with the class label are removed as being uninformative, and of pairs of features that have absolute correlation greater than 0.95 with each other, the feature with the lower correlation with the class label is removed to avoid redundancy. This results in a reduced feature set of 14 features with no loss of discriminative power (cross-validated classification accuracy with all features is within 0.1% of that with reduced features). These 14 features are used in all experiments.

C. Classification Models

The classification module classifies normal activity and malicious activity for the Wireless Sensor Network. The proposed approach applies various methods for machine learning to improve intrusion detection. The classification process uses only the features selected from the WSN-DS dataset. The proposed technique uses Blackhole attacks are identified and categorized using SVM, Decision Tree, Random Forest, Naive Bayes, KNN, Logistic Regression, and XGBoost. Flooding, Grayhole, and Scheduling. In the training stage, the models learn the pattern of network behaviour, while during the testing phase, incoming data is classified based on the learned pattern.

D. Majority-Voting Ensemble

The seven classifiers are ensembled via hard majority voting for each record in the test set, the ensemble predicts the class label that receives the most votes from the seven base classifiers. With an odd-ish number of diverse base learners, there are few ties (less than 0.03% of predictions). The ensemble itself is not trained - it is formed at test time by combining the predictions of the seven base learners. This method doesn't require any additional training data and incurs minimal inference time. The ensemble's diversity is key to its success. Decision Tree and Naive Bayes are strong parametric or axis-aligned assumptions, and hence make different errors in different areas of the feature space to the kernel-based SVM or the distance-weighted KNN. Random Forest and XGBoost are the strongest members of the ensemble but even they do not agree in difficult borderline cases; the ensemble is able to resolve such conflicts by collecting evidence from the whole committee.

E. Evaluation Protocol

The cross-validation will be used to obtain all results. In each fold, 80% of the data will be used for training, including feature selection, normalisation and hyperparameter fitting 20% of the data will be held out for testing. The class proportions will be preserved in both the training and testing data splits. Performance metrics will be provided as the mean from all folds. Overall accuracy will be reported and the precision, recall and F1 score will be calculated using macro-averaged measures. Macro averages do not weigh classes differently thus, classifiers that perform well on the dominant Normal class and poorly on the minority attack.

IV. RESULTS

A. Comparative Performance

Classifier	Accuracy (%)	Precision	Recall	F1-Score
Decision Tree	98.6	0.986	0.985	0.985
Random Forest	99.2	0.992	0.991	0.991
SVM	97.4	0.975	0.973	0.974
Naive Bayes	94.1	0.942	0.940	0.941
KNN	96.3	0.964	0.962	0.963
Logistic Reg.	95.8	0.959	0.957	0.958
XGBoost	**99.4**	**0.994**	**0.993**	**0.994**
Ensemble (MV)	**99.5**	**0.995**	**0.994**	**0.995**

TABLE I. *Cross-Validated Performance of ML Classifiers on WSN-DS (5-fold, macro-averaged)*

B. Feature Space Visualisation

Fig [4.1] Principal Component Analysis (PCA) to reduce the WSN-DS dataset representation to two dimensions clearly shows the class separation of various attacks. The projection into the first two principal components captures the maximum variance, enabling good separation between Normal and DoS attack Flooding, TDMA, Grayhole, Blackhole, etc. The visualization indicates that the Normal class forms a close cluster and is relatively separated from the other classes. On the other hand, TDMA and Flooding attacks share a partial overlap in their features, which may reflect similar distributions and a potential classification difficulty. Nonetheless, we observe moderate separation in the case of Grayhole and Blackhole attacks, suggesting that the chosen features still have discriminatory power for these classes. To sum up, the PCA-based visualization indicating a meaningful structure in the dataset allows for the reasonable application of machine learning algorithms for multi-class intrusion detection with reasonable classification performance.

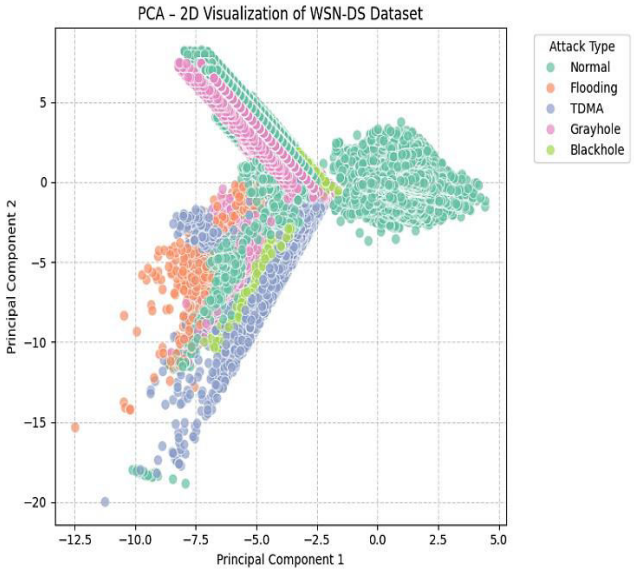

Fig 4.1. PCA-Based 2D Visualization of Class Distribution in WSN-DS Dataset

D. Training Time and Deployment Suitability

The comparative performance of various machine learning approaches is assessed in terms of their training time and classification accuracy, as shown in Fig. [4.2]. After all, it has been seen that XGBoost gives an accuracy of about 96%, but training time is maximum among all the models used. Fitting a random forest model on this dataset gives around 94% accuracy rate with moderate computing cost. In comparison, Support Vector Machine (SVM) gives competitive accuracy (91%), it requires a relatively high amount of training time. Simpler models such as Logistic Regression and KNN provide a balanced trade-off between accuracy and efficiency, achieving accuracy of approximately 90% and 89%, respectively. In the meantime, Decision Tree and Naive Bayes show low accuracy but with low training time, which can be used in low facility conditions. In general, the findings demonstrate that ensemble algorithms surpass standard classifiers in terms of predictive accuracy; though simpler models offer quicker training time with satisfactory performance. However, there is a trade-off between calculation complexity and detection capability.

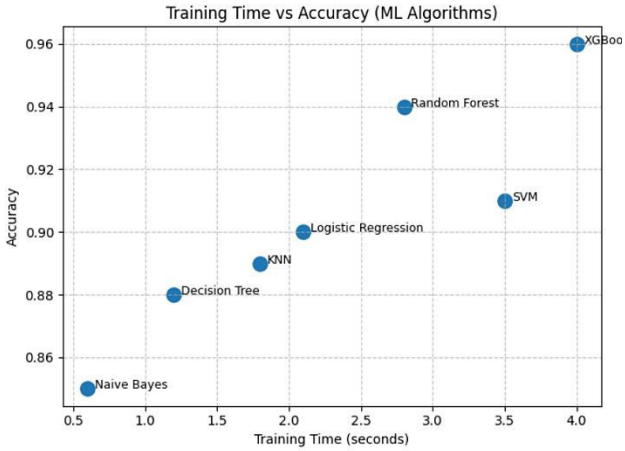

Fig 4.2. Training time vs. Accuracy comparison of machine learning algorithms

V. DISCUSSION

The cross-validated data shown in Table I show several results that have a direct impact on how intrusion detection systems for resource-constrained wireless sensor networks should be designed and implemented. Ensemble techniques utilizing majority-voting classifiers produce the best overall performance with an accuracy of 99.5% and macro-averaged f1 score of 0.995, supporting the established practice of using more than one type of classifier with complementary error characteristics to increase robustness over any one individual classifier, especially for minority class attacks such as grayhole and TDMA, which have highly overlapping feature distributions in PCA space making them particularly difficult to separate from one another. The additional votes from nearest neighbours and logistic regression classes give the majority voting ensemble a particular advantage to classifying these types of attacks that cannot be achieved based upon the results from one of the individual classifiers alone. Among individual classifiers, XGBoost and Random Forest performed well with 99.4% and 99.2% accuracy, respectively; boosting combined with bagging creates the internal diversity needed for these classifiers to be successful. Decision tree performs closely as well with an accuracy of 98.6%, but provides the unique benefit of being able to produce rule-based structures that can be easily audited by a security analyst. Decision trees provide value-added benefits in industrial or regulated environments where explainability is not optional. SVM performed adequately as well with an accuracy of 97.4%, but required approximately three times as much training time as random forest; thus limited its ability to be retrained frequently in order to keep pace with evolving attacks.

VI. CONCLUSION

A study performed a comprehensive evaluation of several classical ML models, using seven classical machine learning classifiers Decision Tree, Random Forest, SVM, Naive Bayes, KNN, Logistic Regression and XGBoost. These classifiers are evaluated for their performance in multi-class intrusion detection for the wireless sensor network. Each of the classifiers was tested on the WSN-DS (wireless sensor networks - Data set) benchmark dataset consisting of 374661 records. The study developed a majority voting ensemble composed of these classifiers using five-fold stratified cross-validation. The overall accuracy of the ensemble was 99.5%, with a macro-averaged F1-particle score of 0.995, which exceeded the performance of any of the individual members of the ensemble, including XGBoost (99.4%, F1: 0.994) and Random Forest (99.2%, F1: 0.991), the top-performing single classifiers. A detailed per-class analysis of the results indicated that two classes of attacks, the Grayhole and Time Division Multiple Access scheduling attacks, were the most difficult to detect due to the overlap of their distribution feature sets in the PCA space. In these borderline cases, an ensemble voting, using the different decision boundaries characterising Naive Bayes, KNN and Logistic Regression, as well as from the tree-based classifiers (Decision Tree, Random Forest), provides a recorded gain in recall of the minority class (i.e. positive recall) over any one of the models. The atmospherics in the computational profiles of each classifier strongly support an operational hierarchy of deployment for each model to ensure the overall security of the wireless sensor network via low-probability detection of malicious activities.

REFERENCES

[1] "A situation-based predictive approach for cybersecurity intrusion detection and prevention using machine learning and deep learning algorithms in wireless sensor networks of Industry 4.0," *IEEE Access*, 2024.

[2] M. Kurtkoti, B. S. Premananda, and K. V. Reddy, "Performance analysis of machine learning algorithms in detecting and mitigating black and grey hole attacks," in *Innovative Data Communication Technologies and Application*. Springer, 2022, pp. 933–943.

[3] M. Sadeghi and H. Alzubaidi, "Fortifying wireless sensor networks using SVM for advanced intrusion detection and attack prevention," *Iraqi Journal of Data Science*, vol. 2, no. 2, pp. 1–13, 2025.

[4] U. Ahmed *et al.*, "Signature-based intrusion detection using machine learning and deep learning approaches empowered with fuzzy clustering," *Scientific Reports*, vol. 15, no. 1, p. 1726, 2025.

[5] A.-G. Mari, D. Zinca, and V. Dobrota, "Development of a machine-learning intrusion detection system and testing of its performance using a generative adversarial network," *Sensors*, vol. 23, no. 3, p. 1315, 2023.

[6] M. B. Musthafa *et al.*, "Optimizing IoT intrusion detection using balanced class distribution, feature selection, and ensemble

machine learning techniques," *Sensors*, vol. 24, no. 13, p. 4293, 2024.

[7] B. Al-Fuhaidi *et al.*, "Anomaly-based intrusion detection system in wireless sensor networks using machine learning algorithms," *Applied Computational Intelligence and Soft Computing*, vol. 2024, pp. 1–19, 2024.

[8] G. Gebremariam, J. Panda, and S. Indu, "Design of advance intrusion detection systems based on hybrid machine learning techniques in hierarchically wireless sensor networks," *Connection Science*, vol. 35, no. 1, 2023.

[9] S. Jadhav *et al.*, "Network intrusion detection system using machine learning," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, pp. 74–81, 2022.

[10] "Machine learning based intrusion detection system," *International Research Journal of Modernization in Engineering Technology and Science*, 2023.

[11] K. NandhaKumar, "A hybrid adaptive development algorithm and machine learning based method for intrusion detection and prevention system," *Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 5, pp. 1226–1236, 2021.

[12] S. Patil *et al.*, "Explainable artificial intelligence for intrusion detection system," *Electronics*, vol. 11, no. 19, p. 3079, 2022.

[13] J. F. C. Garcia and G. E. T. Blandon, "A deep learning-based intrusion detection and prevention system for detecting and preventing denial-of-service attacks," *IEEE Access*, vol. 10, pp. 83043–83060, 2022.

[14] M. Pradeepa and R. Ponnusamy, "Enhancing wireless sensor network security with intrusion detection using recurrent neural networks and optimization algorithms," in *Proc. Int. Conf. Cognitive Robotics and Intelligent Systems (ICC-ROBINS)*, IEEE, 2025, pp. 170–176.

[15] A. V. Turukmane and R. Devendiran, "M-MultiSVM: An efficient feature selection-assisted network intrusion detection system using machine learning," *Computers & Security*, vol. 137, p. 103587, 2024.

[16] T. Zhukabayeva *et al.*, "A traffic analysis and node categorization-aware machine learning-integrated framework for cybersecurity intrusion detection and prevention of WSNs in smart grids," *IEEE Access*, vol. 12, pp. 91715–91733, 2024.

[17] A. R. Ugale and A. D. Potgantwar, "Anomaly-based intrusion detection through efficient machine learning model," *International Journal of Electrical and Electronics Research*, vol. 11, no. 2, pp. 616–622, 2023.

[18] N. Nabi *et al.*, "Machine learning-based anomaly detection for cyber threat prevention," *Journal of Primeasia*, vol. 6, no. 1, pp. 1–8, 2025.

[19] B. B. Jayasingh and M. R. Patra, "A software agent-based approach for fraud detection in network crimes," in *Proc. Asian Applied Computing Conf. (AACC)*, LNCS 3285, Springer, 2004, pp. 310–316.

[20] B. B. Jayasingh and M. R. Patra, "Rule-based evidence mining for network attack," in *Proc. 10th Int. Conf. Information Technology (ICIT)*, IEEE, 2007, pp. 23–28.

[21] V. Chang, S. Boddu, Q. A. Xu, and L. M. T. Doan, "Intrusion detection and prevention with ML algorithms," *International Journal of Grid and Utility Computing*, vol. 14, no. 6, pp. 617–631, 2023.

[22] M. H. Behiry and M. Aly, "Cyberattack detection in WSNs using hybrid feature reduction with AI and ML methods," *Journal of Big Data*, vol. 11, no. 1, p. 16, 2024.

[23] M. A. Talukder *et al.*, "MLSTL-WSN: ML-based intrusion detection using SMOTETomek in WSNs," *International Journal of Information Security*, vol. 23, no. 3, pp. 2139–2158, 2024.

[24] H. M. Saleh, H. Marouane, and A. Fakhfakh, "SGD intrusion detection for a WSN attack detection system using ML," *IEEE Access*, vol. 12, pp. 3825–3836, 2024.

[25] A. John *et al.*, "Cluster-based WSN framework for DoS attack detection based on variable selection ensemble ML," *Intelligent Systems with Applications*, vol. 22, p. 200381, 2024.