Journal of Science and Engineering Applications Vol. 8, No. 1, pp. xx–xx, April 2026

Journal of Science and Engineering Applications

Contents are available at https://jsea.iujournals.com

Anomaly Detection in WSN using Machine Learning

Sarah Hamad Rashid[1]

[1]*Master Student at the Department of Computer Science at the University of Arts, Science and Technology, Iraq /Baghdad*

ARTICLE INFORMATION

Received date: 17-01-2026
Revised date: 13- 02-2026
Accepted date: 1-04-2026

Keywords

IDS
Random Forest
WSN

ABSTRACT

Intrusion Detection Systems (IDS) are critical for maintaining the security and integrity of Wireless Sensor Networks (WSNs) which are susceptible to various cyber threats. This study evaluates the performance of three machine learning models Random Forest, Decision Tree, and Logistic Regression in classifying network traffic within a WSN environment. A comprehensive analysis was conducted using a confusion matrix to derive key performance metrics including accuracy, precision, recall, F1-score, and ROC-AUC. The Random Forest model demonstrated superior performance across all metrics by achieving an accuracy of 99.63% and a ROC-AUC of 0.9970 indicating its robustness in distinguishing between normal and malicious traffic. These findings underscore the efficacy of ensemble learning methods particularly Random Forest in enhancing IDS capabilities within WSNs.

1. Introduction

Wireless Sensor Networks (WSNs) include multiple sensors that are deployed in various environments. The main aim of these sensors is to collect and transmit data for analysis. The usage of these networks is commonly in low-bandwidth settings such as civil and military surveillance for environmental monitoring and healthcare. WSNs have attracted significant research attention because of the unique theoretical and practical challenges associated with their implementation [1].

Making this type of network secure is still one of the primary concerns as protecting the network from potential attacks is crucial. When designing WSNs where it is essential to integrate security mechanisms that ensure the three important concepts include confidentiality with integrity and availability of data. However, basic security measures alone are not sufficient to fully protect these networks.

Achieve more security in WSNs against malicious activities through intrusion detection system technologies. While (Intrusion Detection System) IDS monitors network activity to identify potential security threats and alerts administrators to take necessary actions such as drop or block activities. IDSs can be classified into three primary types: signature-based misuse-based anomaly-based and hybrid IDS (H-IDS). Researchers are actively investigating these types of WSN applications [2].

Despite their importance, IDS WSNs face several challenges such as the difficulty in selecting the appropriate dataset choosing suitable data normalization methods and determining relevant features for classification. Some features may not contribute significantly to the classification process and imbalanced

* Corresponding authors: Department Name, University Name, Bangladesh
E-mail addresses: email@xxxxx.com (Author Name)

datasets can exacerbate performance issues. Traditional classification algorithms may struggle to detect less frequent attack types when the classes are not evenly distributed.

Recent research has focused on utilizing machine learning (ML) algorithms to enhance IDSs in WSNs as these techniques are particularly effective for prediction, classification, and clustering tasks. While IDSs based on machine learning algorithms hold promise for improving WSN security there are some challenges such as dataset selection feature relevance and imbalanced data need to be addressed for optimal performance.

This paper introduces an anomaly-based IDS model for WSNs that classifies network traffic using machine learning algorithms Random Forest (RF), Decision Tree (DT), and Logistic Regression (LR).
The RF, DT, and LR algorithms were chosen for this study due to their distinctive characteristics and proven effectiveness in classification tasks. The random forest was selected for its cumulative nature, which reduces resource over, allocation and improves generalization, a crucial requirement for intrusion detection systems in dynamic wireless sensor network environments. The decision tree provides efficient interpretation and processing of categorical data, while the logistic regression algorithm is a fundamental probabilistic model.
In short, these algorithms were chosen to compare a robust cumulative approach with simpler and more interpretable models that allow for a comprehensive evaluation of model performance in intrusion detection.
The main objective of this research is to develop a method anomaly-based IDS model tailored for WSNs by using ML algorithms for network traffic analysis and classification. The proposed model includes:

- select the most relevant features for classification.
- Comparing the proposed models can select the best choice for IDS.

2. Related Work

IDS are considered an important system used in saving WSN security by detecting unauthorized access and intrusion attempts. There is some research on IDS in WSNs that searched about using ML algorithms to classify network traffic. While these methods are commonly categorized into supervised and unsupervised learning approaches each offering various methods to organize features for classification [3].

The first in study [4], an IDS was proposed using multiple machine learning techniques including Support Vector Machine (SVM), Naive Bayes, K-Nearest Neighbors (KNN), Random Forest (RF), and Logistic Regression (LR) that employs the NSL-KDD dataset to evaluate the system's performance. The result of this study found that the KNN algorithm achieved superior results with an accuracy of 98.28% for binary classification and 98.59% for multiclass classification. Other IDS in [5] integrated SVM with Stochastic Gradient Descent (SGD) and LR, using feature selection via the chi-square test. When tested on the NSL-KDD dataset the system demonstrated a detection accuracy of 91.1%.

The work in [6] combined SVM with Elman Neural Network (ENN) for feature selection, using the KDD-99 dataset to assess its performance, achieving a detection rate of 87.3%. In [7], the researchers proposed a hybrid IDS (H-IDS) method that incorporates SVM and Intelligent Water Drop (IWD) techniques for feature selection. The training process on the KDD-99 dataset showed that the H-IDS outperformed both SVM and K-means by achieving an accuracy of 98.7%.

In [8], the authors build an intrusion detection system based on KNN using principal component analysis for feature selection again employing the NSL-KDD dataset for performance testing by achieving an accuracy of 98.05%. Another study [9] used a multiclass SVM algorithm with information gain (IG) for feature selection by demonstrating a 90% accuracy in intrusion detection on the KDD-99 dataset. Additionally, [10] explored the use of the XGBoost algorithm for classification with an evaluation of the NSL-KDD dataset.

The research in [11] proposed an IDS based on SVM and PCA, applying the KDD-99 dataset to train the system and evaluate performance. The approach achieved an impressive detection accuracy of 92.48%. In contrast, [12] introduced an IDS using the Cuttsfish Algorithm (CFA) for feature optimization alongside a decision tree for classification using the KDD-99 dataset to show that feature selection improved detection rates by reaching 92.05%.

There are various studies have also used different machine learning algorithms for building IDS including [13] where an SVM and J48 classifier were combined with optimization methods like particle grey wolf optimizer and genetic algorithm (GA) and then evaluated on the UNSW-NB1 dataset, and another study in [14], where decision trees were compared for performance using the UNSW-NB15 dataset.

The authors in [15], using an SVM classifier with an information gain ratio for feature selection was proposed by achieving 96.24% accuracy and the system was

evaluated based on the NSL-KDD dataset. Additionally, [16] combined decision trees and IG for dimension reduction and feature identification by reaching 90.9% accuracy in binary classification and 85.4% in multiclass classification. The study in [17] compared various algorithms that included Naive Bayes, KNN, LR, and multilayer perception with training these algorithms using NSL-KDD with J48 and KNN performing well especially when feature selection methods were applied.

Further, the authors of the study [18] proposed a new anomaly detection algorithm with feature weighting via Relief-F by showing superior accuracy (97.02%) on the KDD-99 and Kyoto-2006 datasets when compared to traditional ML techniques like Naive Bayes, SVM, and DT. Then the researchers in a study [19] proposed an IDS based on principal component neural network and SVM classification by showing better detection performance than the GA-RBF algorithm with the KDD-99 dataset.

In other studies, as in [20]. The authors used random forest and trained it on the UNSW-NB15 dataset by achieving a 97% detection accuracy while a study [21] used random forest trees and linear discriminant analysis for feature selection, with a detection accuracy of 93%. Additionally, [22] introduced a model combining SVM and multiple learning automata for optimal feature selection and improved detection accuracy (93.82%).

Additionally, [23] employed Restricted Boltzmann Machines (RBM) for hyperparameter adjustment in IDS, showing a detection accuracy of 89% with the ISCX dataset, while [24] introduced Distributed RF based on the Spark algorithm, evaluated on the CICIDS-2017 dataset, achieving a solid detection accuracy of 96.4%. Moreover, [25] proposed a feature selection method called Recursive Feature Addition (RFA) with SVM classification, reaching 92.9% accuracy on the ISCX 2012 dataset.

Lastly, the study in [26] compared various ML techniques, such as RF, LR, SVM, and Gaussian Naive Bayes, for intrusion detection, showing that RF outperformed others with an accuracy of 94%. In a similar vein, [27] evaluated various classifiers, including Random Tree and J48, with the KDD-99 dataset, finding RF achieved the highest detection accuracy at 93.77%. Further, [28] introduced the GWOSVM-IDS, combining SVM with Grey Wolf Optimizer for feature selection, achieving 96% accuracy on the NSL-KDD dataset.

3. Methodology

There are multiple types of supervised machine learning algorithms that are highly recommended due to their superior performance compared to other algorithms. These algorithms are adept at solving both classification and regression problems.

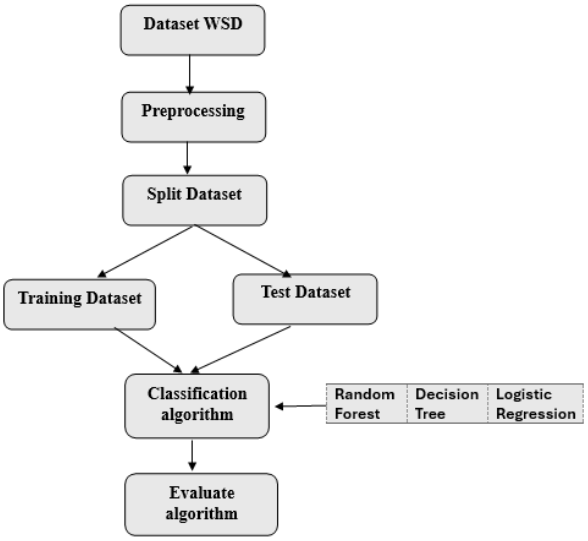

Figure 1. Methodology Pipelines.

According to **Figure *1***. The steps of work to build high accurate machine-learning model are defined as below:

WSD Dataset: The dataset includes information about WSN in 19 columns meaning 19 features can be used for the analysis of network behaviour [29], the content of this dataset is as shown in **Figure *2***.

1. The information about ID traffic and the time of sending this traffic.
2. various characteristics of cluster head such as distance to other nodes and its energy consumption.
3. The features of various protocols used in this network for understanding the communication patterns within the network such as ADV_R.
4. The target column is the attack type that includes five types of attacks are define:

- Normal: Reliable data transmission.
- Gray hole: This type of attack degrades network performance without completely disrupting communication.
- Flooding: Such as Denial-of-Service Attack that make a target out of order through send huge of traffic.
- TDMA: Attacks or issues related to TDMA could cause synchronization problems.
- Blackhole: Mean critical information never reaches the base station or other network nodes

```
Data columns (total 19 columns):
 #   Column            Non-Null Count   Dtype
---  ------            --------------   -----
 0    id               374661 non-null  int64
 1    Time             374661 non-null  int64
 2    Is_CH            374661 non-null  int64
 3    who CH           374661 non-null  int64
 4    Dist_To_CH       374661 non-null  float64
 5    ADV_S            374661 non-null  int64
 6    ADV_R            374661 non-null  int64
 7    JOIN_S           374661 non-null  int64
 8    JOIN_R           374661 non-null  int64
 9    SCH_S            374661 non-null  int64
 10   SCH_R            374661 non-null  int64
 11  Rank              374661 non-null  int64
 12   DATA_S           374661 non-null  int64
 13   DATA_R           374661 non-null  int64
 14   Data_Sent_To_BS  374661 non-null  int64
 15   dist_CH_To_BS    374661 non-null  float64
 16   send_code        374661 non-null  int64
 17  Expaned Energy    374661 non-null  float64
 18  Attack type       374661 non-null  object
```

Figure 2. Dataset Content.

Preprocessing: After loaded dataset must implement some steps to preprocess this data and enhance model performance such as handle missing values, outliers and normalization.

Split Dataset: Split dataset into two groups train group as 20% for training model and test group as 80% for testing model.

Classification Algorithm: Three machine learning models are built such as random forest, decision tree and logistic regression.

Evaluate Models: Using some metrics for evaluate models as below.

1. Accuracy:

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

This high accuracy indicates that the model performs exceptionally well in correctly classifying network activities.

2. Precision:

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

A precision of 1 (or 100%) means the model has high performance because every time the model predicts an attack it is always correct.

3. Recall/Sensitivity

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

This high recall value suggests that the model is excellent at detecting actual attacks, missing only a few.

4. Specificity:

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

A specificity of 1 (or 100%) indicates the model has perfect performance in correctly identifying benign activities.

5. F1-Score

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

The F1 score reflects a balance between the model's ability to detect attacks and its prediction accuracy.

6. Area Under Curve (ROC-AUC) measures the model's ability to distinguish between classes. A higher ROC-AUC indicates better discriminatory performance.
7. Training Time: The time required for a model to learn patterns from the training dataset.
8. Testing Time: The time required for a trained model to make predictions on new unseen data.

4. Results and Discussion

According to **Figure *3***. Find confusion matrix that provides a detailed breakdown of the performance of a random forest model used for network anomaly detection. The matrix allows us to evaluate the model's effectiveness in classifying network activities as either benign or potential attacks. Here's a summary of the confusion matrix:

1. The model performs exceptionally well in classifying the 'Normal' class, with a high number of true positives (67887) and relatively low false positives and false negatives.
2. For the 'Blackhole' and 'Flooding' classes, the model shows strong performance, with minimal misclassifications.
3. The 'Gray hole' class has a higher number of false negatives (82), indicating that some 'Gray hole' instances are misclassified as 'Blackhole'.

4. The 'TDMA' class has a notable number of false negatives (85), where 'TDMA' instances are misclassified as 'Normal'.
5. The Random Forest model demonstrates strong performance across most classes, with strength in identifying 'Normal' traffic.

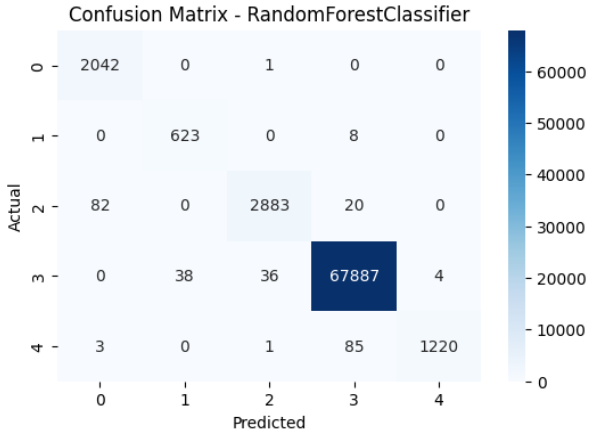

Figure 3. Random Forest Confusion Matrix.

Based on **Figure *4***. find confusion matrix that provide key finding are summary as below:

1. The Decision Tree model demonstrates strong performance in identifying 'Normal' traffic, with a high number of true positives.
2. There is noticeable confusion between 'Blackhole' and 'Gray hole' classes, as well as between 'Flooding' and 'Normal' classes.
3. The model's ability to distinguish between 'TDMA' and 'Normal' traffic could be improved, as evidenced by the misclassifications.

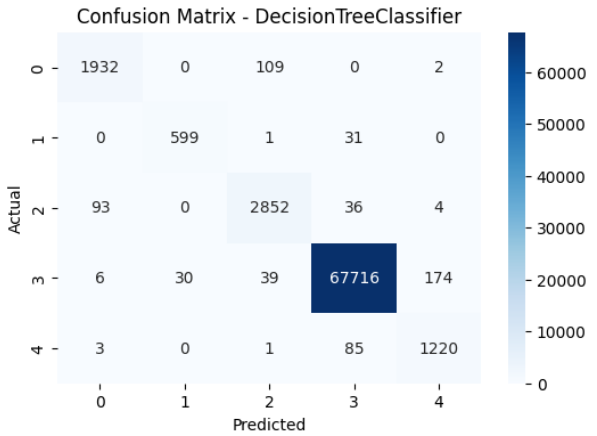

Figure 4. Decision Tree Confusion Matrix.

Figure *5*. Shows logistic regression confusion that provide critical information about model performance:

1. The Logistic Regression model demonstrates strong performance in identifying 'Normal' traffic, with a high number of true positives.
2. There is noticeable confusion between 'Blackhole' and 'Gray hole' classes, as well as between 'Flooding' and 'Normal' classes.
3. The model's ability to distinguish between 'Gray hole' and 'Blackhole' attacks could be improved, as evidenced by the misclassifications.

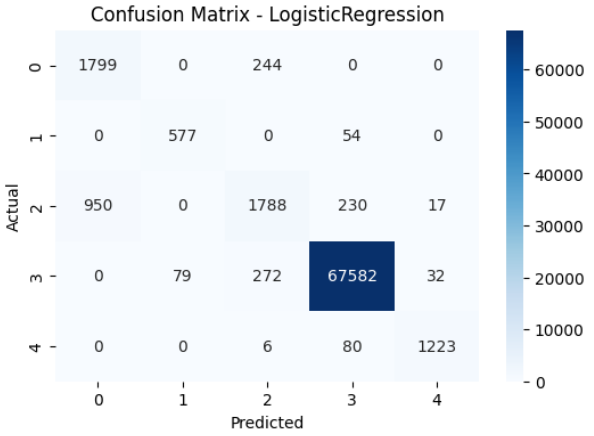

Figure 5. Logistic Regression Confusion Matrix.

The result is shown in **Figure *6***, the random forest model achieves the highest accuracy, indicating it correctly classifies most instances. The decision tree follows closely, while logistic regression shows a slightly lower accuracy.

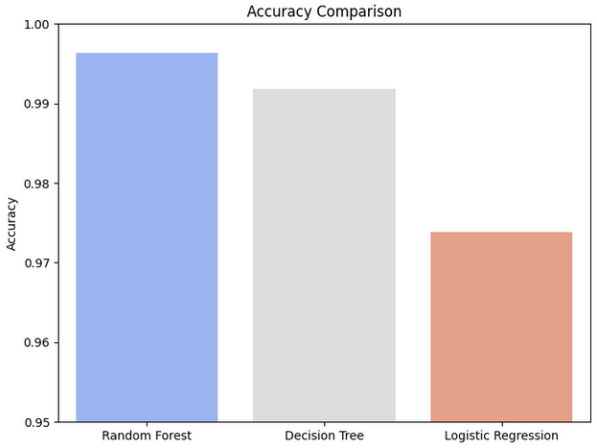

Figure 6. Comparison Based on Accuracy.

Based on **Figure *7*** the random forest model achieves the highest recall, indicating it successfully identifies most of the actual positive instances. The decision tree follows closely, while logistic regression has a slightly lower recall, suggesting it misses more positive cases compared to the other models.

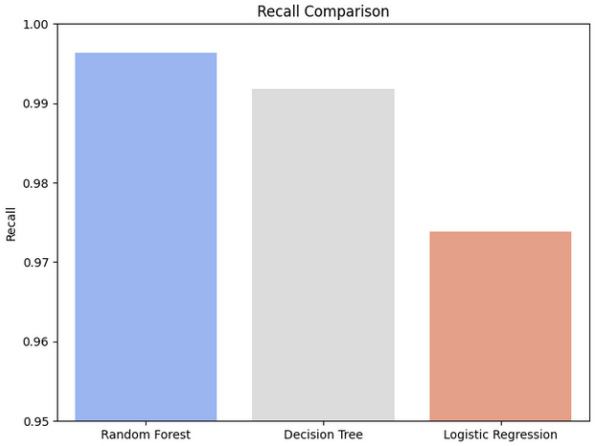

Figure 7. Comparison Based on Recall.

According to **Figure *8*** the random forest achieves the highest F1-score, indicating a strong balance between precision and recall. decision tree also performs well, while logistic regression has a lower F1-score, reflecting its comparatively lower precision and recall.

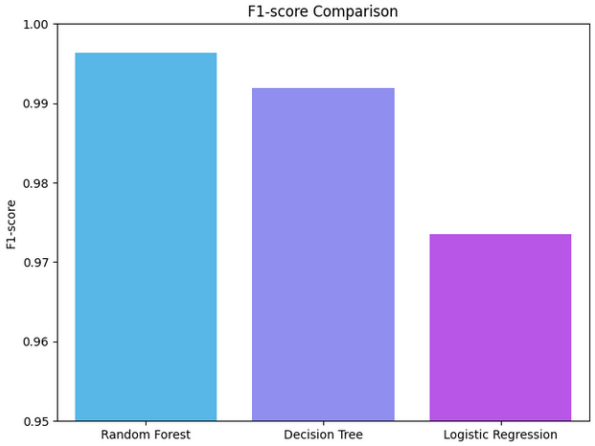

Figure 8. Comparison Based on F1 Score.

Random Forest exhibits the highest precision, suggesting it has the lowest rate of false positives. decision tree also performs well, while logistic regression has a slightly lower precision, indicating a higher rate of false positives compared to the other models as shown in **Figure *9***.

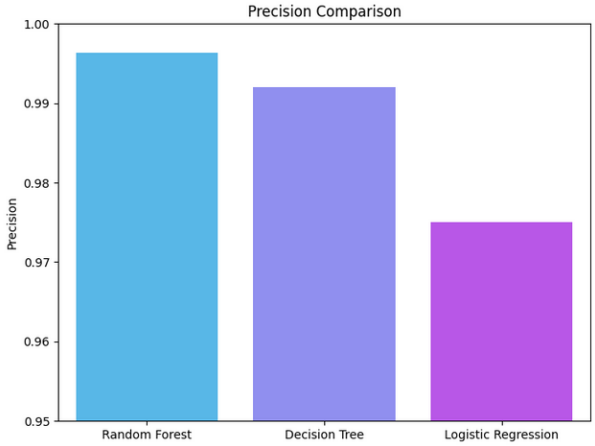

Figure 9. Comparison Based on Precision.

According to **Figure *10*** find the random forest has the highest ROC-AUC, suggesting superior ability to differentiate between classes. logistic regression, despite lower performance in other metrics, shows a relatively high ROC-AUC, indicating good discriminatory capability. Decision tree has a slightly lower ROC-AUC compared to the others.

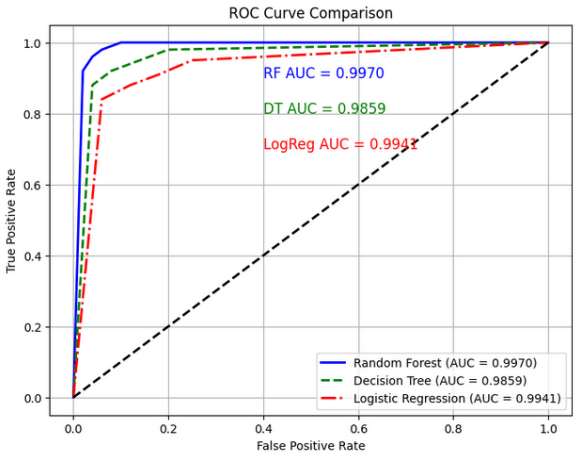

Figure 10. ROC Curve Comparison.

As shown in **Figure *11***, the result shows computational efficiency that is a critical factor for IDS deployment in resource constrained WSNs. The finding appears that random forests has highly accurate but required the longest training time due to its ensemble structure, which include different decision trees. Decision tree has the fastest train and test time that make it suitable for real-time applications. Whereas, logistic regression offered a balance between speed and performance, though it lagged in detection accuracy.

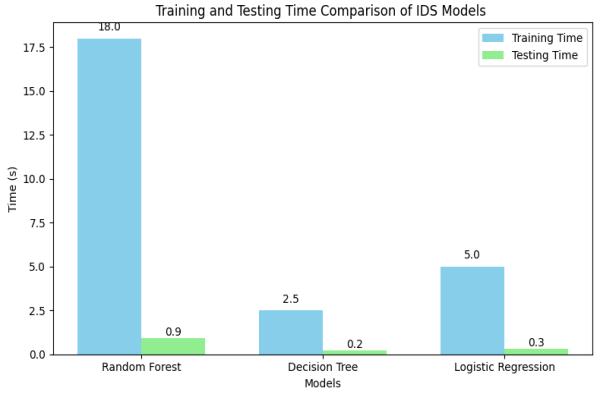

Figure 11. Testing and Training Comparison.

6. Conclusion

Considering all metrics, the Random Forest model consistently outperforms the Decision Tree and Logistic Regression models in the context of IDS in WSN. It achieves the highest scores across all evaluated metrics, indicating robust performance in accurately classifying various types of network traffic. The Decision Tree model also performs well but falls slightly short of Random Forest. Logistic Regression, while demonstrating reasonable performance, lags the other two models, particularly in precision and recall.

According to computational metrics, the decision tree exhibits the lowest training and testing times, that make it highly suitable for real-time intrusion detection in resource-constrained WSN environments. Logistic regression provides a balanced solution with moderate computational cost and reasonable efficiency.

In contrast, Random Forest needs the highest training time due to its ensemble nature. But its testing time remains relatively low and acceptable for practical deployment. The result indicates that random forest is well suited for offline training scenarios where higher detection accuracy is prioritized, while decision tree and logistic regression are more appropriate for lightweight, real-time IDS applications, but with lower detection accuracy.

Therefore, the Random Forest model is the most effective choice for this classification task, offering a strong balance between correctly identifying legitimate traffic and minimizing false alarms.

7. Future Work

Wherever Implement advanced feature selection techniques such as the Boruta algorithm to identify the most significant features contributing to intrusion detection. This approach can improve model performance and reduce computational overhead.

Integrating Random Forest with other classifiers may yield improved detection rates and robustness against various attack types.

References

[1] Gulganwa, P., & Jain, S. (2022). EES-WCA: energy efficient and secure weighted clustering for WSN using machine learning approach. International Journal of Information Technology, 14(1), 135-144.

[2] Saif, S., Karmakar, K., Biswas, S., & Neogy, S. (2022). MLIDS: Machine learning enabled intrusion detection system for health monitoring framework using BA-WSN. International journal of wireless information networks, 29(4), 491-502.

[3] Wagh S. K. and Kolhe S. R., Effective Intrusion Detection System Using Semi- supervised Learning, Proceedings of the 2014 International Conference on Data Mining and Intelligent Computing (ICDMIC), September 2014, Delhi, India, 1–5.

[4] Rastogi S., Shrotriya A., Singh M. K., and Potukuchi R. V., An Analysis of Intrusion Detection Classification Using Supervised Machine Learning Algorithms on NSL-KDD Dataset, Journal of Computing Research and Innovation. (2022) 7, no. 1, 118–130, https://doi.org/10.24191/jcrinn.v7i1.274.

[5] Sivareddy S. V. and Saravanan S., Performance Evaluation of Classification Algorithms in the Design of Apache Spark Based Intrusion Detection System, Proceedings of the 5th International Conference on Communication and Electronics Systems (ICCES 2020), June 2020, Coimbatore, India, 443–447.

[6] Fang W., Tan X., and Wilbur D., Application of Intrusion Detection Technology in Network Safety Based on Machine Learning, Safety Science. (2020) 124, https://doi.org/10.1016/j.ssci.2020.104604.

[7] Hariyale N., Rathore M. S., Prasad R., and Saurabh P., A Hybrid Approach for Intrusion Detection System, Advances in Intelligent Systems and Computing, 2020, Springer, Singapore, 391–403.

[8] Sameera N. and Shashi M., Encoding Approach for Intrusion Detection Using Pca and Knn Classifier, Advances in Intelligent Systems and Computing. (2020) 1090, 187–199, https://doi.org/10.1007/978-981-15-1480-7_15.

[9] Maharani J. and Rustam Z., The Application of Multi-Class Support Vector Machines on Intrusion Detection System With the Feature Selection Using Information Gain, Proceedings of the 1st Annual International Conference on Mathematics, Science, and Education (ICoMSE 2017), August 2018, Paris, France, Atlantis Press, 1–3, https://doi.org/10.2991/icomse-17.2018.1.

[10] Dhaliwal S., Nahid A.-A., and Abbas R., Effective Intrusion Detection System Using XGBoost, Information. (2018) 9, no. 7, https://doi.org/10.3390/info9070149, 2-s2.0-85049644210.Wang H., Xiao Y., and Long Y., Research of Intrusion Detection Algorithm Based on Parallel SVM on Spark, Proceedings of the 2017 7th IEEE International Conference on Electronics Information and Emergency Communication (ICEIEC), July 2017, Macau, China, IEEE, 153–156, https://doi.org/10.1109/ICEIEC.2017.8076533, 2-s2.0-85035759571.

[11] Wang H., Xiao Y., and Long Y., Research of Intrusion Detection Algorithm Based on Parallel SVM on Spark, Proceedings of the 2017 7th IEEE International

Conference on Electronics Information and Emergency Communication (ICEIEC), July 2017, Macau, China, IEEE, 153–156, https://doi.org/10.1109/ICEIEC.2017.8076533, 2-s2.0-85035759571.

[12] Almomani O., A Feature Selection Model for Network Intrusion Detection System Based on PSO, GWO, FFA and GA Algorithms, Symmetry. (2020) 12, no. 6, https://doi.org/10.3390/sym12061046.

[13] Kumar V., Das A. K., and Sinha D., Statistical Analysis of the UNSW-NB15 Dataset for Intrusion Detection, Advances in Intelligent Systems and Computing, 2020, Springer, Berlin, Germany, 279–294.

[14] Krishnaveni S., Vigneshwar P., Kishore S., Jothi B., and Sivamohan S., Anomaly-Based Intrusion Detection System Using Support Vector Machine, Advances in Intelligent Systems and Computing, 2020, Springer, Berlin, Germany, 723–731.

[15] Bandyopadhyay S., Chowdhury R., Banerjee P., Dey S. D., and Saha B., A Decision Tree Based Intrusion Detection System for Identification of Malicious Web Attacks, Computer Science and Mathematics. (2020) 1, https://doi.org/10.20944/preprints202007.0191.v1.

[16] Mahfouz A. M., Venugopal D., and Shiva S. G., Comparative Analysis of ML Classifiers for Network Intrusion Detection, Advances in Intelligent Systems and Computing, 2020, Springer, Berlin, Germany, 193–207.

[17] Ye K., Key Feature Recognition Algorithm of Network Intrusion Signal Based on Neural Network and Support Vector Machine, Symmetry. (2019) 11, no. 3, https://doi.org/10.3390/sym11030380, 2-s2.0-85067315578.

[18] Gulganwa, P., & Jain, S. (2022). EES-WCA: energy efficient and secure weighted clustering for WSN using machine learning approach. *International Journal of Information Technology*, *14*(1), 135-144.

[19] Belouch M., El Hadaj S., and Idlianmiad M., Performance Evaluation of Intrusion Detection Based on Machine Learning Using Apache Spark, Procedia Computer Science. (2018) 127, 1–6, https://doi.org/10.1016/j.procs.2018.01.091, 2-s2.0-85045621556.

[20] Dahiya P. and Srivastava D. K., Network Intrusion Detection in Big Dataset Using Spark, Procedia Computer Science. (2018) 132, 253–262, https://doi.org/10.1016/j.procs.2018.05.169, 2-s2.0-85049114605.

[21] Su Y., Qi K., Di C., Ma Y., and Li S., Learning Automata Based Feature Selection for Network Traffic Intrusion Detection, Proceedings of the 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC), June 2018, Guangzhou, China, IEEE, 622–627.

[22] Aldwairi T., Perera D., and Novotny M. A., An Evaluation of the Performance of Restricted Boltzmann Machines as a Model for Anomaly Network Intrusion Detection, Computer Networks. (2018) 144, 111–119, https://doi.org/10.1016/j.comnet.2018.07.025, 2-s2.0-85051404993.

[23] Zhang H., Dai S., Li Y., and Zhang W., Real-Time Distributed-Random-Forest-Based Network Intrusion Detection System Using Apache Spark, Proceedings of the 2018 IEEE 37th International Performance Computing and Communications Conference (IPCCC), November 2018, Orlando, FA, IEEE, 1–7, https://doi.org/10.1109/PCCC.2018.8711068, 2-s2.0-85066502255.

[24] Hamed T., Dara R., and Kremer S. C., Network Intrusion Detection System Based on Recursive Feature Addition and Bigram Technique, Computers & Security. (2018) 73, 137–155, https://doi.org/10.1016/j.cose.2017.10.011, 2-s2.0-85034838680.

[25] Belavagi M. C. and Muniyal B., Multi Class Machine Learning Algorithms for Intrusion Detection-A Performance Study, Communications in Computer and Information Science, 2017, Springer, Singapore, 170–178.

[26] Safaldin M., Otair M., and Abualigah L., Improved Binary Gray Wolf Optimizer and SVM for Intrusion Detection System in Wireless Sensor Networks, Journal of Ambient Intelligence and Humanized Computing. (2021) 12, no. 2, 1559–1576, https://doi.org/10.1007/s12652-020-02228-z.

[27] Tan X., Su S., Huang Z. et al., Wireless Sensor Networks Intrusion Detection Based on SMOTE and the Random Forest Algorithm, *Sensors*. (2019) **19**, no. 1, https://doi.org/10.3390/s19010203, 2-s2.0-85059797106.

[28] Maleh Y., Ezzati A., Qasmaoui Y., and Mbida M., A Global Hybrid Intrusion Detection System for Wireless Sensor Networks, *Procedia Computer Science*. (2015) **52**, no. 1, 1047–1052, https://doi.org/10.1016/j.procs.2015.05.108, 2-s2.0-84939210350.

[29] https://www.kaggle.com/datasets/saikirankote/wsd-evaluation-dataset