Journal of Engineering Science and Technology
Vol. 21, No. 2 (2026) 422 - 435

AN ENHANCED ENSEMBLE LEARNING MODEL-BASED IOT INTRUSION DETECTION SYSTEM TO IMPROVE AIOT AGAINST CYBER ATTACKS

SHAYMA WAIL NOURILDEAN*,
NASSREN NAJM ABD ALWAHED, YOUSRA ABD MOHAMMED

University of Technology- Iraq
*Corresponding Author: Shayma.w.nourildean@uotechnology.edu.iq

Abstract

The spread of Internet of Things (IoT) has revolutionized many fields by enabling extensive connection options and data -driven decisions. IoT network expansion had introduced significant security weaknesses, which required strong security mechanisms. Integration of artificial intelligence (AI) into IoT, called artificial intelligence of things, provides increased ability to detect and attenuate the intelligent danger. This study presents an ensemble-based Intrusion Detection System (IDS) utilizing a soft voice technique to improve the detection of the cyber attackers in the IoT network. CIC-IOT2023 and IOTID20 data sets were evaluated by empirical evaluation, including ransomware and various types of cyber-attack. The proposed ensemble model achieved better performance with the accuracy rate of 95.06% and 99,998% respectively on CIC-IOT 2023 and IOTID20 data sets, with high precision, recall, F1-score and ROC-AUC values indicating better performance. Comparative analysis demonstrated that the proposed model improves individual models and improves the reliability of safety systems and efficiency. These findings suggest that ensemble learning represents a promising opportunity to detect future infiltration in the smart environment.

Keywords: Accuracy, AIoT, IDS, Machine learning, Ransomware.

1. Introduction

IoT is a network of physical entities which combine sensors, software and technology via Internet, transmits data with other systems [1]. Artificial Intelligence (AI) had been incorporated into the IoT system to introduce an Artificial intelligence of things (AIoT) to be an effective way to process and store large data in IoT [2]. It is a new study area used in different domains and can get many benefits. IoT acts as a digital nervous system, while artificial intelligence acts as the brain [3].

IoT-based applications are the targets of cyber-economic attacks. Such attacks damage not only human life, but also physical welfare and natural environment. IoT-identification of new attacks in IoT-based smart environment requires an effective IDS. Designed to monitor hosts or networks for security breaches, IDSs alert the administrator on discovery [4, 5].

An IDS is a software program or physical device which detect harmful network activity and trigger an alarm. This helps to determine the source of an assault, therefore enabling the chance of destroying it. Furthermore, it facilitates the detection of attacks such as denial of service (DoS) and man-in-the-middle (MitM) [6, 7]. Data mining, deep learning and machine learning are elements of artificial intelligence that emulate the human brain's data processing capabilities and are nowadays extensively utilized to enhance the quality and accuracy of IDSs [8, 9].

This paper aimed to build an efficient IDS for IoT system against a number of cyber-attacks defined in CIC-IoT2023 and IoTID20 including (Ransomware, DoS, DDoS, Mirai, Benign, MiTM, Recon-OSScan, DNS_Spoofing, Recon, Brute-Force) [10] and IoTID20 dataset for binary and multiple classification. The proposed ensemble model is examined against some existing traditional machine learning models for a number of performance metrics like (Fl-score, Recall, Accuracy and precision). In this paper, the area under the curve (ROC-AUC) is also determined for the proposed model against other models to show its effectiveness. To guide this study, the following research questions are formulated:

- How effective is the proposed ensemble learning model (Decision Tree, Random Forest, XGBoost with soft voting) compared to traditional machine learning models in detecting diverse IoT cyberattacks across CIC-IoT2023 and IoTID20 datasets?
- Can the ensemble model improve both binary and multiclass classification performance in the IoT intrusion detection with respect to precision, accuracy, F1-score, recall, and ROC-AUC?
- Does the integration of ensemble learning enhance the robustness and reliability of IDSs in AIoT environments, thereby contributing to stronger security in smart ecosystems?

2. Related Work

The development of IDSs employing various Machine and Deep Learning techniques has been the focal point of research targeted at addressing security challenges in IoT networks [7]. Dini, et al. [11] examine IDS based machine

learning techniques in terms of different algorithms and performance metrics on three datasets: KDD 99, CSE-CIC-IDS 2018 and UNSW NB15. Different machine learning approaches were assessed. Results indicate that Decision Tree (DT) and Random Forest (RF) surpass the other models with accuracy rates of 99.4% and 100% for multiclass and binary classification respectively. Dhumal and Pingale [12] conducted an examination of (IDS) using deep learning methodologies, with a particular focus on the Convolutional Neural Network (CNN) algorithm with CICIDS 2018 dataset. The fundamental objective is to improve the accuracy and dependability of detecting cyber-attacks in an IoT networks. Xu et al. [13] provides an effective IDS for safeguarding IoT networks through the utilization of XGBoost classifiers using BGWO and RFE-XGBoost with two step feature selection techniques to discern essential feature subsets. The results indicated that the proposed strategy surpasses the other methods across five datasets with 1.0 accuracy on the WUSTL IIOT-2021 and BoT IoT datasets.

A full stack classification model developed in to increase the classification matrix needed to sprinkle the classification matrix as a contingent modelling and throwing function choice choices [14]. The proposed model using the 'TON IOT dataset' usually improved precision, accuracy, recalling, F1 score of individual models indicating its benefit as a measure of IoT network safety problems. Learning-based optimization of intrusion detection system (TLBO-IIDS) developed by Kaushik et al. [15] in this system, an IoT protects the network from intrusion with lower costs. After an intensive analysis, TLBO-AIDS improved the modern algorithm, performed better than 40% from the BAT algorithm and 40% from the genetic algorithm (GA). Alsulami et al. [1] presented an IP-IDS, an improved Intrusion Detection System (IDS) that can recognize MQTT-referenced inputs in a variety of datasets. Several prior ML/DL models trained on similar data sets were outperformed by the proposed model. In recent years, a thorough assessment of machine learning algorithms for IoT network intrusion detection was presented by Baich et al. [16]. With a minimum forecast time of 0.4 seconds and an accuracy of 99.26%, the data analysis shows that the Fisher score performed the best in decision trees.

While existing works have shown promising results in improving IDS performance for IoT networks, they also present notable limitations. Most previous work is based heavily on a standalone machine learning or deep learning model (e.g., CNN, RF, XGBoost), which can achieve high accuracy on certain datasets but often fails to generalize across various types of attacks and real-world scenarios. Moreover, certain methods are based on specific data sets like CICIDS2018 or UNSW-NB15 that do not completely represent the complexity of emerging IoT attacks and threats such as ransomware or hybrid ones, despite being applied in wide range of studies. Feature selection methods like RFE and BGWO have also yielded better detection accuracy in some studies; however, they introduce additional computational overhead limiting the scalability. Conversely, some previous works have claimed near-perfect accuracy (although comparisons between the binary and multi-class setting are rarely conducted rigorously) under different conditions.

Instead, the contribution of this paper is to construct an ensemble IDS which combines Decision Tree, Random Forests and XGBoost using a soft voting scheme. The ensemble differs from the single classifier models since it tends to utilize various classifiers in a complementary way, hoping that they will compensate each other. Additionally, in this work, the model is tested on two challenging and diverse IoT datasets (CIC-IoT2023 and IoTID20) for both binary

and multi-class classification tasks which show its wider applicability compared to many previous works. By incorporating the ROC-AUC analysis, the model is validated to be more robust and reliable by considering if a packet captured with successive time does discriminate benign/malicious better. Thus, this work offers the negative answer for the open question whether decentralized IDS is feasible and necessitates broad SCIP infrastructure.

3. Theoretical Concepts

The combination of IoT and AI has been used in several commercial areas as a catalyst for future changes, which has been pressed by progress in both AI and IoT technology [17].

3.1. AIoT

The modern term AIoT (Artificial Intelligence of Things) combines two important ideas: IoT and AI. In an IoT, an efficient and intelligent data processing is necessary to make the best use of the information produced by this data. AI may be used to analyse and use the data for decision-making or problem-solving. Without artificial intelligence, the Internet of Things would only be partially functional. Furthermore, the growing popularity of modern AIoT systems and applications is associated with additional difficulties, including complexity, efficiency, scalability, accuracy, and resilience [18, 19]. IoT devices integrated with artificial intelligence (AI) can analyse data, make choices, and execute actions autonomously, without human intervention [20].

Artificial intelligence (AI), deep learning (DL) and machine learning (ML) always refer to intelligent software or systems [21].

3.2. IoT-cyber security

The availability, confidentiality, and integrity are universally acknowledged concepts in cybersecurity. Numerous assaults from diverse external or internal sources are disclosed within the IoT network [22]. Applications of the IoT are susceptible to cyber-kinetic attacks. This type of attack poses a risk to human life, physical health, or the environment. Cyber-kinetic attacks on critical infrastructures reliant on the Internet of Things are often complex. Various diverse methods and techniques are employed in the performance of these jobs. Daily, fresh cyberattacks are initiated, making it exceedingly difficult to avoid all of them. The intrusion detection prevention processes are both crucial to reducing the effects of cyberattacks. The forms of cyber threats in IoT include [23]: Botnets, Brute force, Man-in-the-middle, Credential assaults, Denial of service, Eavesdropping, Firmware attacks, Privilege escalation, and Ransomware..

3.3. Intrusion detection systems

IoT intrusion is an illegal activity or behaviour affecting the IoT system. An attack affecting the integrity, accessibility and damage to the information is classified as an intrusion [24]. The phrase "intrusion detection" refers to the identity of harmful activities carried out against the information system [25]. In addition to identifying potential threats and protecting the network from malicious attacks and illegal access, IDS acts as a defensive obstacle. In order to combat network intrusion and other

hazards, IDSs are important components used in modern computer network infrastructure. They monitor the IoT network and identify any illegal intrusion. When internal and external threats are detected, ID generates flags or issuance alerts [26].

IDS is required to monitor network activity and detect possible breaches of security [11]. In order to be reliable, safe and profitable for the Internet of Thing services, it is important to detect real-time intrusion on IoT units [27]. Network-IDS (NID) and host-IDS (HIDS) are two types of intrusion detection systems. Depending on their identity method, IDS can be divided into two types: anomaly-based IDSs and signature-based IDSs [9].

In order to detect malicious activity in the IoT network, different types of machine learning (ML) and Deep Learning (DL) techniques have been developed using IDS.

3.4. Machine learning

Figure 1 illustrated the positioning of ML and DL within AI [21].

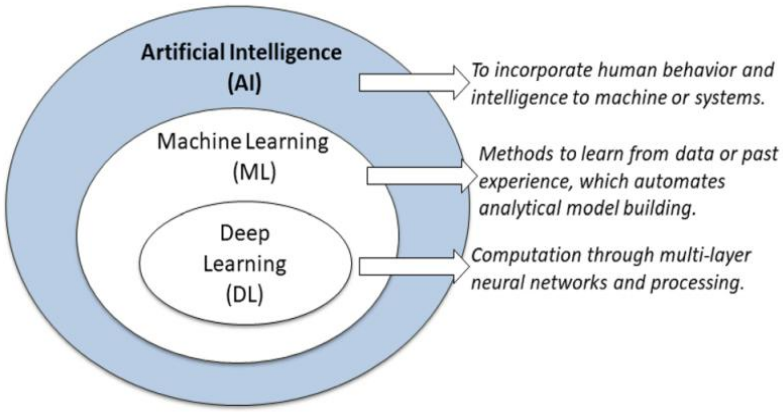

Fig. 1. ML and DL [21].

Machine learning algorithms provide a clear edge of traditional recognition techniques. They can handle high quality data, non-led data and identify complex patterns and rules from a wide dataset, providing them more effective inputs for complex systems [13].

4. Research Method

This paper aimed to develop an enhanced IDS in an IoT systems that can combat various cyber threats identified in CIC-IoT2023 and IoTID20 for both binary and multiclass classification. CIC-IoT2023 [28] with Ransomware datasets offers a larger number of devices, diverse usage profiles, and both IP and non-IP device data, providing valuable insights for developing effective detection methods. IoTID20 Dataset is an open-source IEEE dataset that allows us to detect IoT networks' abnormal activity. It serves as a benchmark for identifying anomalous activity across IoT networks and would be a valuable resource for new IDS techniques development in IoT environments. Table 1 shows the dataset description.

Table 1. Dataset description.

Dataset	Dataset Size (Samples)	Number of Features	Number of Classes	Attack Types (Categories)
IoTID20	625,783	86 Total	5	1. **Mirai** (Botnet DDoS) 2. **Scan** (Reconnaissance) 3. **DoS** (Denial-of-Service) 4. **MITM ARP Spoofing** (Man-in-the-Middle) 5. **Normal** (Benign Traffic)
CIC IoT2023	238,687	47	35	DDoS-ICMP_Flood, DDoS-UDP_Flood, DDoS-TCP_Flood, DDoS-PSHACK_Flood, DDoS-SYN_Flood, DDoS-RSTFINFlood, DDoS-SynonymousIP_Flood, DoS-UDP_Flood, DoS-TCP_Flood, DoS-SYN_Flood, BenignTraffic, Mirai-greeth_flood, Mirai-udpplain, Mirai-greip_flood, DDoS-ICMP_Fragmentation, MITM-ArpSpoofing, DDoS-ACK_Fragmentation, DDoS-UDP_Fragmentation, DNS_Spoofing, Recon-HostDiscovery, Recon-OSScan, Recon-PortScan, DoS-HTTP_Flood, VulnerabilityScan, DDoS-HTTP_Flood, DDoS-SlowLoris, DictionaryBruteForce, SqlInjection, BrowserHijacking, CommandInjection, Backdoor_Malware, XSS, Uploading_Attack, Recon-PingSweep
Ransomware Dataset	22,391	85	6	RANSOMWARE_LOCKERPIN, RANSOMWARE_WANNALOCKER, RANSOMWARE_RANSOMBO, RANSOMWARE_KOLER, RANSOMWARE_PORNDROID, RANSOMWARE_PLETOR
Combined Dataset	261,078	55 (after feature extraction)	40	All attack types from both datasets combined

An ensemble model integrates several distinct models to enhance overall predictive accuracy and resilience. This model integrates Decision Tree, Random Forest, and XGBoost into an ensemble employing a soft voting method. Subsequently, it trains the ensemble, assesses its performance through precision, recall, accuracy, F1-score, and visualizes its predictions with a confusion matrix.

This code snippet integrates Decision Tree, Random Forest, and XGBoost models into an ensemble employing a soft voting technique. This model is examined against some traditional machine learning models using multiple assessment measures. (Decision Trees DT [29], Xgboost [30], Logistic Regression LR [31], k-nearest neighbour KNN [32], Naive Bayes NB [33], CATBOOST CAT [34]). It was analysed for binary and multiple categorizations.

4.1. Data preprocessing and feature selection

The experiments were conducted in Python using Scikit-learn, NumPy, and SciPy, with random seeds fixed to ensure reproducibility. For both CIC-IoT2023 and IoTID20 datasets, preprocessing involved removing duplicate rows, replacing inf/-inf with NaN, dropping unresolved missing values, and excluding non-predictive identifiers such as timestamps.

All features were numeric; categorical attributes (rare in these datasets) were one-hot encoded. Scale-sensitive models (e.g., Logistic Regression, KNN) were standardized using Standard Scaler, while tree-based learners (DT, RF, XGBoost) were trained on raw features. To preserve class balance, data was split using a stratified 80:20 train–test ratio, and 5-fold cross-validation was applied on the training set for model tuning.

Feature selection followed a multi-stage process:

- Low-variance filtering to discard near-constant features
- Pruning the Correlation with ($|r| \geq 0.95$) to reduce redundancy.
- Ranking of the most informative predictors based on Mutual Information.
- Recursive Feature Elimination with Cross-Validation (RFECV) with Random Forest (optimized for macro F1-score).

This pipeline resulted in a fixed subset of 55 features that were repeated over all binary and multiple classifications. Crucially, the whole selection process was fitted on training folds only to prevent information leakage. Final evaluation was examined on the test set using Accuracy, Precision, Recall, F1 and ROC-AUC to ensure fair model comparison.

4.2. Algorithm steps

The general stages of the proposed technique are

- Step 1: Apply the preprocessing pipeline and feature selection to determine structure and missingness of data.
- Step 2: Splits the dataset into test and train using split-ratio 80:20, taking the last 20% for testing set and keeping first 80% for training.
- Step 3: Include required libraries.
- Step 4 Create a list to hold the base classifiers that we want to combine.
- Step 5: Add single models (Decision Tree, Random Forest, XGBoost) to the list.
- Step 6: Create a soft voting ensemble to average the class probabilities predicted by each classifier, thus combining the predictions of the three models.
- Step 7: The training data is used to train the ensemble model.

- Step 8: The accuracy i.e. the ratio of correct predictions is measured with test data.
- Step 9: Output Accuracy, Precision, Recall and F1-Score as shown.

- Accuracy: General correctness.
- Reciprocal: The number of relevant elements that have been retrieved.
- Recall: Number of relevant items found.
- F1-Score: The harmonic means of accuracy and recall, balancing both measures.

The empirical analysis of the paper written in Python Programming Language. Machine learning models are built using the Sklearn library. It is based on essential libraries such as NumPy, SciPy or matplotlib. Scikit-learn features various classification, regression and clustering algorithms which can be easily used as a ML library in Python. The quality of the classifiers was evaluated on two pre-processed dataset data. The results are displayed as follows:

- Precision, accuracy, F1-score and recall were measured for the ensemble model compared to a variety of existing models in both binary and multi-class classification withCIC-ID2023 dataset. depicted in Table 2.

Table 2. Evaluation metrics for IoTID2023 dataset +Ransomware attacks.

Machine Learning	Accuracy	Precision	Recall	F1-Score
Ensemble	95.0647%	95.0265%	95.0647%	95.0297%
xgboost	93.6878%	93.5681%	93.6878%	93.5758%
KNN	89.8039%	90.1108%	89.8039%	89.5440%
DT	94.8139%	94.8267%	94.8139%	94.8161%
LR	69.7545%	69.7647%	69.7545%	64.5945%
NB	57.5973%	66.9938%	57.5973%	55.0630%
Catboost	93.0366%	93.0503%	93.0366%	92.8612%

- Accuracy, precision, F1-score and recall were measured for the ensemble model compared to a variety of existing models in both binary and multi-class classification with IoTID20 dataset depicted in Table 3.

Table 3. Performance metrics for IoTID20 dataset.

ML Models	Accuracy	Precision	Recall	FI-Score
Ensemble	99.998%	99.997%	99.998%	99.999%
xgboost	99.990%	99.990%	99.990%	99.990%
KNN	99.773%	99.773%	99.773%	99.773%
DT	99.995%	99.970%	99.967%	99.985%
LR	87.083%	88.170%	87.083%	87.009%
NB	70.824%	89.372%	70.824%	75.967%
Catboost	99.996%	99.996%	99.996%	99.996%

Tables 1 and 2 show that the proposed method can detect the IoT cyber-attacks effectively and stably, with an accuracy of 95.0647% for CIC-IDS2023 dataset and 99.998% for IoTID20, which improves intelligent system reliability. He results highlight that our approach is of advantage over existing IDS schemes in detecting the attacks using multiple DL systems on two various datasets.

- A ROC curve illustrates the True Positive Rate (TPR) in relation to the False Positive Rate (FPR) across different threshold configurations. It facilitates the visualization of the performance of a binary classifier. An effective model will exhibit a curve that closely adheres to the top-left corner (high True Positive Rate, low False Positive Rate). AUC (Area Under the Curve): AUC quantifies the whole two-dimensional area beneath the ROC curve. It is a singular metric that encapsulates the classifier's efficacy in differentiating across classes. As the AUC approaches 1.0, it signifies a perfect classifier. The ROC-AUC was analysed for each model, including the suggested ensemble model for each dataset, as seen in Figs. 2 and 3, respectively.

As illustrated in Figs. 2 and 3, the ROC of this ensemble model is the better top-right corner than existing models. The AUC of this ensemble model is 0.9993 for IoTID20 and 0.95025 for CIC-IoT2023 which are closer to one and thus indicating the better performance.

The superior performance of the ensemble model can be explained by its ability to combine the complementary strengths of Decision Tree, Random Forest, and XGBoost. Each base learner contributes differently:

- Decision Trees (DT) are simple and interpretable, but they tend to overfit when trained on high-dimensional IoT traffic features.
- Random Forests (RF) mitigate overfitting by averaging multiple trees, improving stability but sometimes underestimating rare attack classes.
- XGBoost is powerful at capturing complex, non-linear relationships and handling imbalanced data through gradient boosting, but it can be sensitive to parameter tuning.

By integrating these classifiers through soft voting, the ensemble averages their probability outputs, reducing variance and avoiding the weaknesses of any single model.

For instance, even with the wrong DT generalization of subtle attack behaviors in dataset, RF and XGBoost make up for it by learning more consistent decision boundaries. On the other hand, if RF's majority classes are underpruned - and its minority class is overfitted, XGBoost's boosting iterations can potentially restore those patterns.

This is why the ensemble not only reached better accuracy 95.06%(CIC-IoT2023) and 99.998% (IoTID20)) but also obtained better recall and F1-scores to maintain balanced detection between majority (class of benign data) and minority class (attack data). The fact that the ensemble has higher ROC-AUC values (0.95025) and (0.9993) highlighted a more robust performance in accurately distinguishing between benign and malicious traffic at different thresholds.

In general, the strength of ensemble comes from its resilience against overfitting and better generalization to varied and heterogeneous IoT attacks than stand-alone machine learning models.

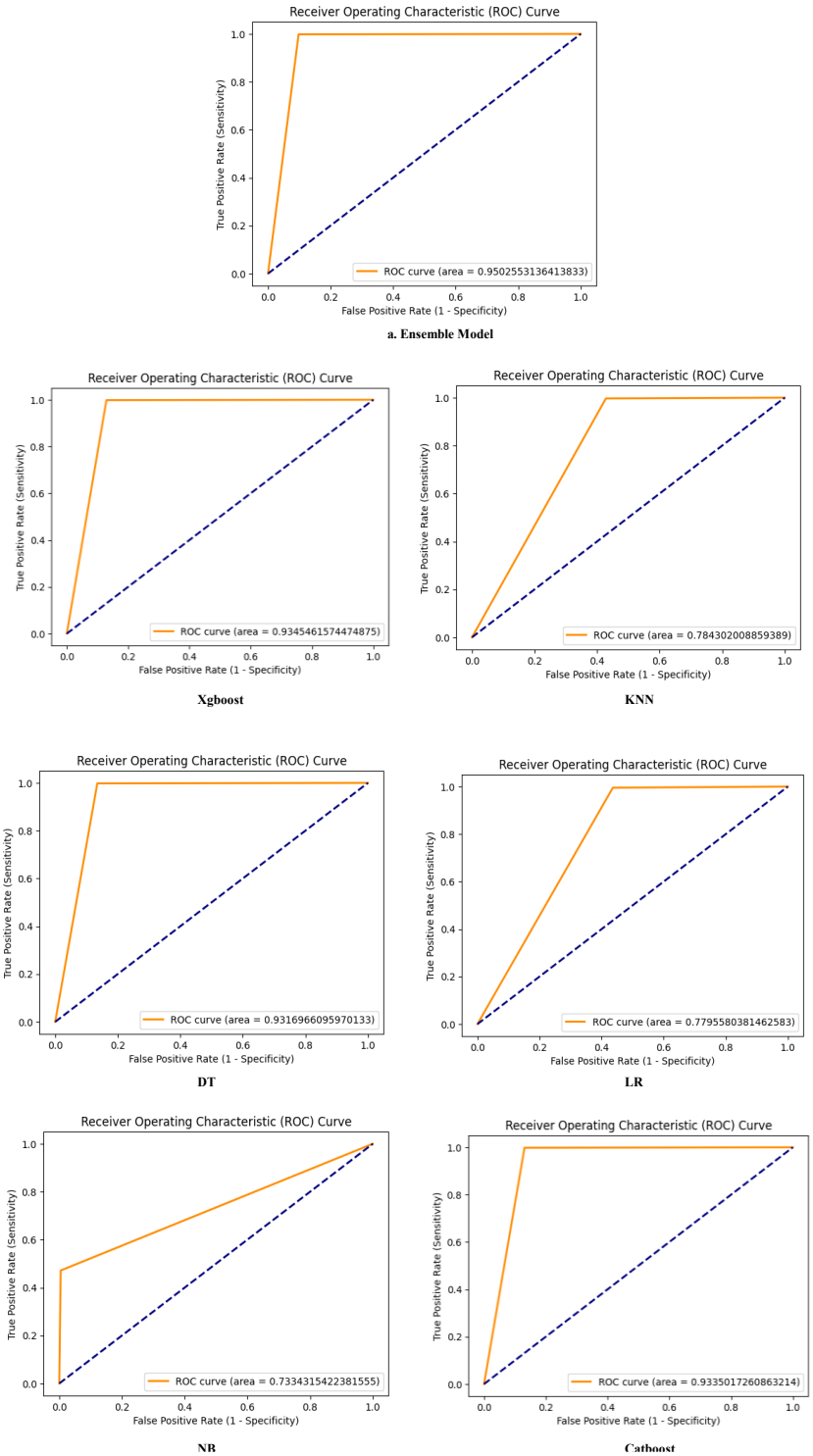

Fig. 2. ROC-AUC (Area Under the Curve) of all the models across for CIC-ID2023 dataset.

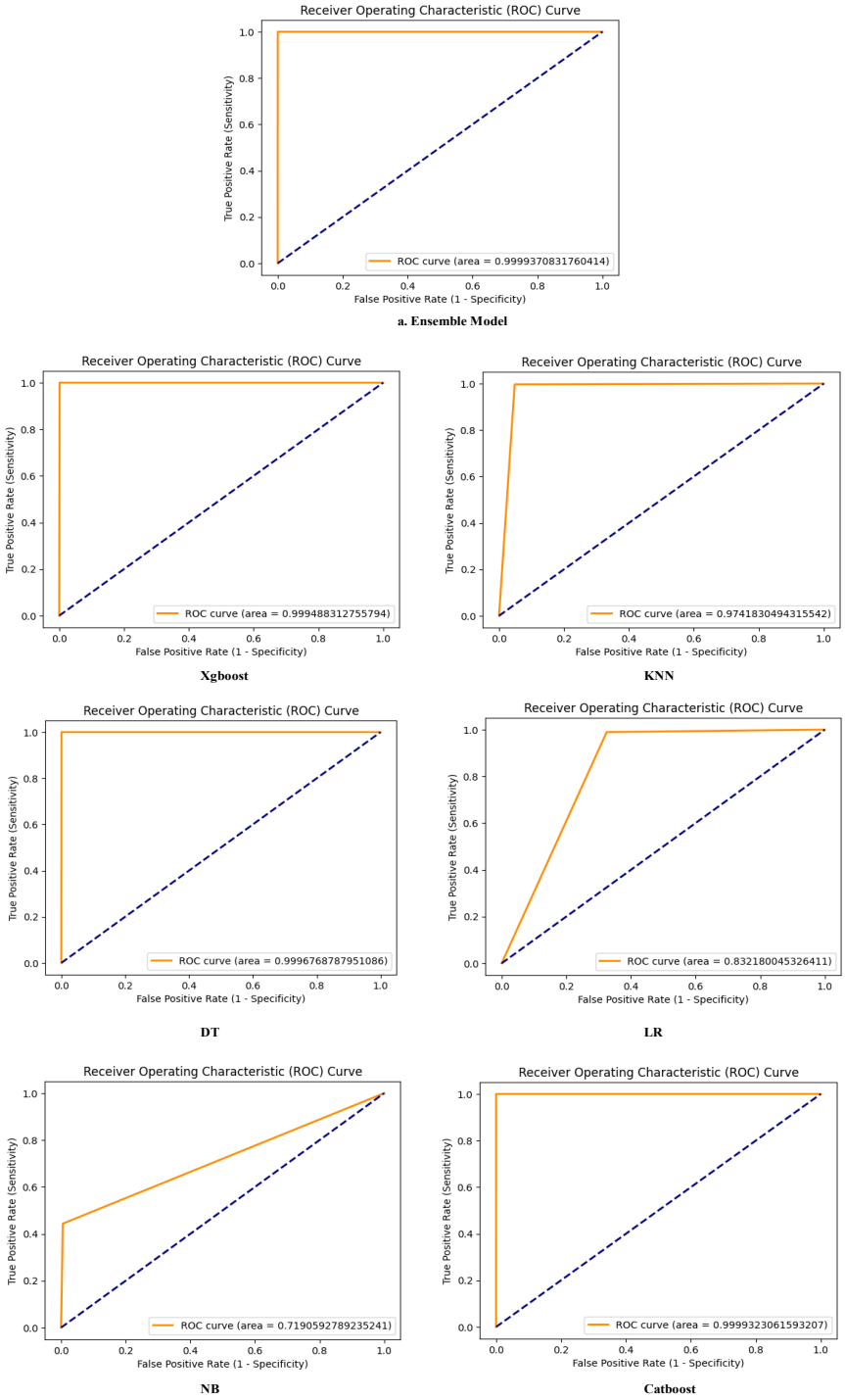

Fig. 3. ROC-AUC (Area Under the Curve) of all the models across for IoTID20 dataset.

5. Conclusion

AI and IoT have been integrated to form the new challenging research domain known as AIoT, which has improved the data processing, decision-making and automation functionalities of IoT networks. As the size of IoT systems grows rapidly, there are also more and more cybersecurity threats attached to them, indicating the urgent development of efficient and effective IDS. In this study, an ensemble model is proposed that combines Decision Trees, Random Forests and XGBoost based on soft voting for enhancing the IoT intrusion detection. It showed better performance of the proposed model compared to the standard machine learning techniques.

The ensemble model in particular attained accuracies of 95.06% and 99.998% on the CIC-IoT2023 and IoTID20 datasets, respectively, with relatively high precision, recall, F1-score, as well as good AUC values indicative of its soundness and reliability. These empirical outcomes confirm the effectiveness of ensemble learning approaches to improve IDSs detection performance for enhancing IoT networks security, leading to hardening smart environments against a broad range of cyberattacks.

For future work, we will consider extending the ensemble approach using adaptive learning techniques and testing on diverse real world IoT deployment scenarios to improve resilience and generalisation. In conclusion, the proposed ensemble framework presents a promising direction for enhancing IoT security and contributes meaningfully toward the realization of resilient, intelligent, and autonomous IoT ecosystems

References

1. Alsulami, R.; Alqarni, B.; Alshomrani, R.; Mashat, F.; and Gazdar, T. (2023). IoT protocol-enabled IDS based on machine learning. *Engineering, Technology & Applied Science Research*, 13(6,) 12373-12380.
2. Nourildean, S.W. (2024). Artificial internet of things (AIoT): A review. *Proceeding of the International Conference on Electrical Computer and Energy Technologies* (*ICECET*, 2024), Sydney, Australia.
3. Nozari, H.; Szmelter-Jarosz, A.; and Ghahremani-Nahr, J. (2022). Analysis of the challenges of artificial intelligence of things (AIoT) for the smart supply chain (Case s: FMCG Industries). *Sensors*, 22(8), 2931.
4. Alosaimi, S.; and Almutairi, S.M. (2023). An intrusion detection system using BoT-IoT. *Applied Sciences*, 13(9), 5427.
5. Nourildean, S.W.; Mefteh, W.; and Frihida, A.M. (2025)). Hybrid deep learning model-based intrusion detection system to improve artificial internet of things against cyber-attacks. *Proceedings of the International Conference on Advanced Information Networking and Applications*, 270-280, Barcelona, Spain.
6. Vanin, P. et al. (2022). A study of network intrusion detection systems using artificial intelligence/machine learning. *Applied Sciences*, 12(22), 11752.
7. Nourildean, S.W.; Jasim, S.I.; Abdulhadi, M.T.; and Jaber, M.M. (2022). Point coordination mechanism based mobile ad hoc network investigation

against Jammers. *Eastern-European Journal of Enterprise Technologies*, 119(9), 45-53.

8. Hdidou, R.; and Alami, M.E. (2024). Advancements in intrusion detection systems for internet of things: A state-of-the-art and comprehensive analysis of machine learning algorithms. *Journal of Theoretical and Applied Information Technology*, 102(2), 602-618.
9. Sarker, I.H. (2021). Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science*, 2(6) 1-20.
10. Nourildean, S.W.; and Hassib, M.D. (2024). IoT-based MANET performance improvement against jamming attackers in different mobile applications. *e-Prime-Advances in Electrical Engineering*, *Electronics and Energy*, 8, 100615.
11. Dini, P. et al. (2023). Overview on intrusion detection systems design exploiting machine learning for networking cybersecurity. *Applied Sciences*, 13(13), 7507.
12. Dhumal, C.T.; and Pingale, D.S.V. (2024). Analysis of intrusion detection systems: techniques, datasets and research opportunity. *Proceedings of the International Conference on Innovative Computing & Communication* (*ICICC* 2024).
13. Xu, B.; Sun, L.; Mao, X.; Ding, R.; and Liu, C. (2023). IoT intrusion detection system based on machine learning. *Electronics*, 12(20), 4289.
14. Almotairi, A.; Atawneh, S.; Khashan, O.A.; and Khafajah, N.M. (2024). Enhancing intrusion detection in IoT networks using machine learning-based feature selection and ensemble models. *Systems Science & Control Engineering*, 12(1), 2321381.
15. Kaushik, A.; and Al-Raweshidy, H. (2024). A novel intrusion detection system for internet of things devices and data. *Wireless Networks*, 30(1), 285-294.
16. Baich, M.; Hamim, T.; Sael, N.; and Chemlal, Y. (2022). Machine learning for iot based networks intrusion detection: A comparative study. *Procedia Computer Science*, 215, 742-751.
17. S. El-Gendy. (2020). IoT Based AI and its Implementations in Industries. *Proceedings of the* 15*th International Conference on Computer Engineering and Systems* (*ICCES*, 2020), Cairo, Egypt, 1-6.
18. Sung, T.-W.; Tsai, P.-W.; Gaber, T.; and Lee, C.-Y. (2021). Artificial intelligence of things (AIoT) technologies and applications. *Wireless Communications and Mobile Computing*, 2021(1), 9781271.
19. Nourildean, S.W.; Mefteh, W.; and Frihida, A.M. (2025). An artificial intelligence of things intrusion detection framework for mitigating cyber and ransomware threats in IoT networks. *Journal of Soft Computing and Data Mining*, 6(1), 333-346.
20. Nourildean, S.W.; Zaiter, M.J.; Hassib, M.D.; and Mohammed, Y.A. (2025). IoT Based RPL Routing Protocol Improvement against DDoS Cyber Attack Using NetSim v14.1 Simulation Program. *International Journal of Electrical and Electronic Engineering & Telecommunications*, 14(1), 35-42.
21. Sarker, I.H. (2022). AI-based modeling: Techniques, applications and research issues towards automation, intelligent and smart systems. *SN Computer Science*, 3(2), 158.

22. Sicato, J.C.S.; Singh, S.K.; Rathore, S.; and Park, J.H. (2020). A comprehensive analyses of intrusion detection system for IoT environment. *Journal of Information Processing Systems*, 16(4), 975-990.
23. Das, R.; and Gündüz, M.Z. (2021). Analysis of cyber-attacks in IOT-based critical infrastructures. *International Journal of Information Security Science*, 8(4), 122-133.
24. Khraisat, A.; and Alazab, A. (2021). A critical review of intrusion detection systems in the internet of things: techniques, deployment strategy, validation strategy, attacks, public datasets and challenges. *Cybersecurity*, 4(1), 18.
25. Odeh, A.; and Taleb, A.A. (2022). IoT security challenges and intrusion detection systems. *Encyclopedia. pub IoT*, 52541, 1-10.
26. Eric, G.; and Jurcut, A. (2022). Intrusion detection in internet of things systems: A review on design approaches leveraging multi-access edge computing, machine learning, and datasets. *Sensors*, 22(10), 3744.
27. Awajan, A. (2023). A novel deep learning-based intrusion detection system for IoT Networks. *Computers*, 12(2), 34.
28. Nourildean, S.W.; Mefteh, W.; and Frihida, A.M. (2025). DTXG-RF-based Intrusion Detection System for Artificial IoT Cyber Attacks. *Engineering, Technology & Applied Science Research*, 15(1), 19610-19614.
29. Sarailidis, G.; Wagener, T.; and Pianosi, F. (2023). Integrating scientific knowledge into machine learning using interactive decision trees. *Computers & Geosciences*, 170, 105248.
30. Chen, T.; and Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco California, USA, 785-794.
31. Sperandei, S. (2014). Understanding logistic regression analysis. *Biochemia Medica*, 24(1), 12-18.
32. Sun, J.; Du, W.; and Shi, N. (2018). A Survey of kNN Algorithm. *Information Engineering and Applied Computing*, 1(1), 1-10.
33. Wickramasinghe, I.; and Kalutarage, H. (2021). Naive Bayes: Applications, variations and vulnerabilities: a review of literature with code snippets for implementation. *Soft computing*, 25(3), 2277-2293.
34. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; and Gulin, A. (2018). Catboost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*, 31.