Journal of Theoretical and Applied Information Technology
15th April 2026. Vol.104. No.7

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DUAL-THRESHOLD FEATURE SELECTION AND LIGHTWEIGHT ADAPTIVE SOFT-VOTING ENSEMBLE FOR NETWORK INTRUSION DETECTION IN IoT ENVIRONMENT

VEENA S BADIGER[1], GOPAL K SHYAM[2]

[1]Research scholar, School of engineering, Presidency university, Karnataka, India.

[2]Professor, HOD, School of engineering, Presidency university, Karnataka, India.

E-mail: [1]veenasbadi@gmail.com, [2]gopalshyambabu@gmail.com

ABSTRACT

Network intrusion detection systems (NIDS) is an important tool for protecting networking devices. In an resource constrained network environment such as IoT safeguarding these devices is essential. The use of lightweight NIDS is necessary to find intrusion in resource-constrained IoT environments. Achieving high detection accuracy with real-time efficiency is challenging therefore, a novel two-stage lightweight NIDS is developed to improve intrusion detection in IoT networks. The study presents three important novel components. First Dual-Threshold Random Forest Feature Selection (DT-RFFS) methodology decreases the feature space by 68-71 percent and increases the rate of inference by 12-15. Confidence weighted adaptive soft voting (CW-ASV) strategy adapts the weight of the classifiers based on the performance and improves the minority attack detector by 2.3%. Third Confidence Based Early Exit (CBEE) mechanism leaves earlier high confidence benign traffic with a significant reduction in the computational overhead and a high detection accuracy. Proposed architecture performs binary classification in first stage to distinguish benign traffic from malicious network activity. Subsequently, the second stage identifies specific intrusion categories using Soft Voting ensemble learner consisting Decision Tree, Random Forest, LightGBM, XGBoost, and AdaBoost as Base learners. The dataset is balanced using SMOTE. The experiment was conducted on CIC-ToN-IoT and RT-IoT2022 datasets. CIC-ToN-IoT and RT-IoT2022 experiment results achieved the accuracy of 99.8 and 98.6-99.7 in stage 1 and stage 2 respectively, and a precision, recall and F1-score of 99 percent. The log-loss values are between 0.01 and 0.10, and AUC-ROC values are between 0.9 and 1.0 indicating high reliability in probability prediction and overall class differentiation. Cross dataset testing has a generalization accuracy of 90-91%. The experiment demonstrated inferences at low computation price. Model achieved a latency of less than 20 ms, and a throughput of 690 to 720 packets/second, showing its real-time performance and deployment in resource-constrained IoT network conditions.

Keywords - *Cyber-attacks, Ensemble Learning, Intrusion Detection, IoT Security, Lightweight Classifier, Soft Voting.*

1. INTRODUCTION

The Internet of Things (IoT) has become one of the major technological paradigms impacting people's daily lives such as it is applied in smart houses, healthcare, agriculture, economies, and transportation. IoT boosts efficiency and decision making by interconnecting physical devices and digital systems and providing cybercriminals more attack surface. It is important to focus on the vulnerability of network because the traditional security measures including firewalls are not adequate to counterattack of the sophisticated types. Most advanced attacks are detected by the Intrusion Detection Systems (IDS). Thus, it is needed to protect IoT networks in case of unauthorized access and threats.

The main role of the IDS is to identify the malicious activity and facilitate the provision of timely response [1]. It is estimated that there will huge rise in the usage of IoT devices [2], due to this the issue of cybersecurity is also going to rise [3]. The network-based IDS (NIDS) is deployed at the networking nodes [4] which analyze the network traffic in order to search any anomalous traffic.

Usage of machine learning (ML) algorithms to

create intrusion detection systems (IDS) has become popular. Many ML based IDS have been suggested but single algorithm based approaches have drawbacks that include low detection rates against a wide range of attacks, high false positive rates and overfitting, making it less likely to generalize to unknown traffic.

Developing an effective Network intrusion detection system (NIDS) is essential in resource constrained environments, where computing power, memory and energy are limited. Standard NIDS systems based on machine learning and deep learning are beneficial in most conventical IT environments, but become more complicated to implement because of high computation costs and long inference times. These models have difficulties in satisfying security issues such as huge volumes of data and problems of real-time detection and limited processing power.

To overcome these challenges, in the current research an ensemble learning has been applied, where multiple ML algorithms are been combined to improve detection accuracy and robustness. Ensemble models leverage complementary strengths of classifiers, reducing individual weaknesses.

In this work, a novel lightweight two-stage Network Intrusion Detection System (NIDS) integrating three key innovations is proposed, namely: (i) Dual-Threshold Random Forest Feature Selection (DT-RFFS) for adaptive dimensionality reduction, (ii) Confidence-Weighted Adaptive Soft Voting (CW-ASV) for dynamically assigning classifier weights based on prediction confidence and historical performance, and (iii) Confidence-Based Early Exit (CBEE) to reduce computational overhead by terminating high-confidence predictions at an earlier stage. The proposed framework employs a hierarchical detection strategy in which Stage 1 performs binary classification to distinguish benign and malicious traffic, while Stage 2 identifies specific attack categories using an ensemble of Decision Tree, Random Forest, LightGBM, XGBoost, and AdaBoost classifiers. The framework is evaluated on two heterogeneous IoT datasets, CIC-ToN-IoT and RT-IoT2022, and demonstrates superior performance in terms of accuracy, precision, recall, F1-score, AUC-ROC, kappa score, and log loss compared to individual classifiers.

The key contributions of this work are:

- Novel Two-Stage Hierarchical Architecture: A lightweight intrusion detection framework that sequentially performs binary and multiclass classification, reducing computational overhead by 34% through a Confidence-Based Early Exit (CBEE) mechanism.
- Dual-Threshold Random Forest Feature Selection (DT-RFFS): An enhanced feature selection technique that reduces feature dimensionality by 68-71% while improving inference speed by 12-15% through adaptive thresholding.
- Confidence-Weighted Adaptive Soft Voting (CW-ASV): A dynamic weighting strategy that adjusts classifier influence based on per-class historical performance, improving minority attack detection by 2.3%.
- Comprehensive Resource Benchmarking: First study to provide detailed energy consumption, memory usage, and inference time comparisons against deep learning models on actual IoT hardware (Raspberry Pi 4).
- Cross-Dataset Generalization Analysis: Systematic evaluation of model transferability across heterogeneous IoT datasets, demonstrating that mixed-source training achieves 99.5% accuracy on unseen data.
- Multi-Metric Robustness Validation: Extensive evaluation using accuracy, precision, recall, F1-score, AUC-ROC, log loss, and kappa score, achieving log loss as low as 0.01, confirming exceptional model reliability.

The primary objective of this study is to design and evaluate a lightweight network intrusion detection framework suitable for resource-constrained IoT environments. To achieve this goal, the following specific objectives are defined:

(i) To improve intrusion detection performance in both binary and multiclass scenarios.

(ii) To reduce feature dimensionality and computational overhead using the proposed DT-RFFS method.

(iii) To enhance ensemble reliability through confidence-weighted adaptive soft voting.

(iv) To minimize inference latency using a

confidence-based early exit mechanism.

(v) To validate the generalization capability of the proposed framework across heterogeneous IoT datasets.

The remainder of the paper is structured as follows, section 2 summarizes the existing works and their limitations. Section 3 explains the proposed two stage network intrusion detection system. Section 4 presents the experimental setup and compares the results with state-of-the-art works. Section 5 explains the limitations of the proposed study. Finally, Section 6 concludes the paper and explains the future work.

2. REVIEW WORK

In the IoT domain, Intrusion detection system based on machine learning, deep learning, artificial intelligence and federated learning has been proposed to detect threats. Various studies have proved improved accuracy, reduced false positives, or enhanced robustness, still challenges exist. Some of the works are been summarized in this section.

Ensemble based IDS have shown promising results. Rohini et al. [5] proposed a weighted majority voting ensemble. SMOTE was applied for data balancing and Arithmetic Optimization Algorithm with Butterfly Optimization Algorithm to enhance machine learning classifiers. Model was not evaluated for error handling. Zouhri et al. [6] investigated univariate and multivariate feature selection filters with classifiers such as MLP, SVM, XGBoost, and RF across multiple datasets.

Fatima et al. [7] proposed ELIDS. Ensemble feature selection selects robust features to generate lightweight IDS to identify DDoS. Empirical benefits were asserted in the work by the reduction of feature set. ELIDs was aimed at lightweight IDS in the IoT although the paper fails to tell much about energy consumption, or latency. The model is developed to detect DDoS attack.

Niktabe [8] proposed an LSTM profiling model with SHAP based XAI for global and local explanations. The work was tested on IoT network in order to detect botnet and DDoS attacks. XAI layer makes the model more transparent and is suitable for actionable threat analysis. Limitations of the work was it lacks of latency and throughput metrics measurement.

Musthafa et al. [9] proposed a technique for IoT intrusion detection which combines class balancing using SMOTE feature selection using statistical techniques. Ensemble machine learning algorithms by SVM bagging and LSTM stacking was used to get better detection performance. The model exhibited high accuracies on the publicly available datasets with good AUC. Work Focuses on binary class classification and does not report multiclass classification.

Soni et al. [10] developed binary IDS on IoT based on ensemble algorithms such as XGBoost and LightGBM and evaluated on publicly available dataset. They found increased accuracy in comparison to previous methods. The study is limited to the binary classification.

Odeh et al. [11], proposed ensemble deep learning architecture consisting of models such as CNN, LSTM and GRU, based on voting ensemble for IoT intrusion detection. Model achieved accuracies around 99.7% and F1-scores of 99.8% for their best models CNN-LSTM and CNN-GRU respectively. Model provides limited knowledge about the deployment in resource constrained IoT device.

From this literature review, several areas of improvement can be identified:

- Although existing studies report higher accuracy and lower false alarm rates, they do not provide other important evaluation metrics such as log loss, and kappa score. As a result, the reported performance of existing IDS may not accurately reflect the model's robustness or generalization.
- Most existing IDS approaches primarily focus on DDoS detection, and are not designed or evaluated for identifying a broader range of attack types.
- Many studies rely on older or limited datasets, which may not adequately represent modern IoT traffic patterns or emerging attack categories.
- Several techniques claim to be lightweight, yet they lack detailed evaluation of resource efficiency metrics such as latency, energy consumption, or memory usage, and often omit multiclass performance metrics such as AUC-ROC and log loss.
- Existing models demonstrate limited generalization capabilities and have not been validated on heterogeneous or real-world IoT

environments, reducing their deployment feasibility.

- Therefore, there remains a need for a robust, adaptable, and resource efficient IDS capable of detecting multiple attack types while maintaining low computational overhead in constrained IoT environments.

The current research proposes a novel lightweight two-stage network intrusion detection system using an enhanced soft voting ensemble machine learning approach to detect cyber-attacks in resource-constrained IoT environments. To address limitations in existing works, the proposed model introduces three key innovations: (1) Dual-Threshold Random Forest Feature Selection (DT-RFFS) that reduces feature dimensionality by 68-71% while improving inference speed by 12-15%, (2) Confidence-Weighted Adaptive Soft Voting (CW-ASV) that dynamically adjusts classifier weights based on per-class historical performance, improving minority attack detection by 2.3%, and (3) Confidence-Based Early Exit (CBEE) mechanism that reduces computational overhead by 34% for normal traffic without compromising detection accuracy.

The model is considered lightweight due to the strategic selection of base classifiers that achieve an optimal balance between high performance and computational efficiency. It combines low-complexity yet effective algorithms including Decision Tree (DT), Random Forest (RF), XGBoost (XGB), AdaBoost (ADB), and Light Gradient Boosting Machine (LGBM), which are known for their fast training and inference capabilities. The two-stage hierarchical structure first performs binary class classification on the network traffic as attack and non-attack by using CW-ASV. Next identifies the type of attacks that are recognized when the network traffic is identified as attack traffic by using CW-ASV of the multi-class classification. This two-stage hierarchical structure is more efficient than the other existing methods as types of classifications are trained on the entire data set. This reduces the computational time of the classification training, also reduces the computing cost.

To test model effectiveness, extensive analysis was performed with the help of various performance measures such as accuracy, precision, recall, F1-score, AUC-ROC, kappa score, and log loss. The model proposed demonstrated binary classification accuracy (99.8) and multiclass classification accuracy (98.6-99.7) on both CIC-ToN-IoT and RT-IoT2022. The log loss of 0.1- 0.01 is an excellent error management and prediction accuracy. Cross-dataset validation demonstrated 90-91% generalization accuracy with single dataset training, and enhanced to 99.5% with combined heterogeneous data training, 71-83% reduced energy consumption versus deep learning alternatives, and a latency of less than 20ms and a throughput of 690-720 packets/second, indicated the appropriateness of the model in real-time deployment in resource-constrained IoT applications.

3. PROPOSED METHODOLOGY

This paper proposes a novel Lightweight Network Intrusion Detection System (LNIDS) named iLNIDS using enhanced soft voting ensemble learning approach for two-stage cyber-attack detection in resource-constrained IoT systems. There are three contributions to this paper: (1) Dual-Threshold Random Forest Feature Selection (DT-RFFS), (2) Confidence-Weighted Adaptive Soft Voting (CW-ASV), and (3) Confidence-Based Early Exit (CBEE) mechanism. The motivation and problem analysis are elaborated in Section 3.1 and the proposed model is illustrated in Section 3.2. 3.3 explains datapreprocessing and 3.4 explains dataset balancing. Training and Testing is explained in 3.6. The proposed feature selection Dual Threshold Random Forest Selection (DT-RFFS) is explained in 3.7. The detailed explaination of proposed Confidence -Weighted Adaptive Soft Voting (CW-ASV) for classification is explained in 3.8. Summary of contributions of propsoed work is explained in 3.9 and 3.10 sumarizes on the performance metrics used for experiment.

3.1 Problem Analysis And Motivation

Most existing research in Intrusion Detection Systems has focused on performing binary class classification [22, 23], which fails to provide individual attack categorization necessary for targeted mitigation. Methods designed to conduct multiclass classification are able to detect the precise type of intrusion but they also pose a very large computational burden, since all data is processed against all classes at once [18]. Other methods execute binary and multiclass classification separately [24,25], a factor that adds to the cost of computations because redundant operations are performed and resources are not utilized efficiently.

3.2 Study Design and Experimental Protocol

In order to evaluate how effective the proposed lightweight intrusion detection framework for resource-constrained IoT environments, the research was conducted using a quantitative experimental design. The experimental workflow consists of data preprocessing, feature selection, model training and performance evaluation. The CIC-ToN-IoT and RT-IoT2022 publicly available datasets were used for the development of the IoT security framework to provide heterogeneous and realistic representations of IoT traffic.

The preprocessed data was given to the proposed Dual-Threshold Random Forest Feature Selection (DT-RFFS) for feature reduction. The reduced features have been used to train the two-step detection method by using proposed framework.

In order to establish a valid experimental protocol, 80:20 dataset splitting for training and testing was applied. Stratified sampling was used to evaluate the model's generalization ability, by performing cross-dataset validation and tested the model trained on one dataset against another.

The study further evaluated the computational efficiency of the framework using performance metrics.

3.3 Overview Of Proposed Framework

The proposed framework uses the combination of five base classifiers, such as Decision Tree (DT), Random Forest (RF), XGBoost (XGB), AdaBoost (ADB), and Light Gradient Boosting Machine (LGBM). These algorithms were chosen strategically because they are known to provide a good performance and computation efficiency thus they are suitable in resource constrained environments.

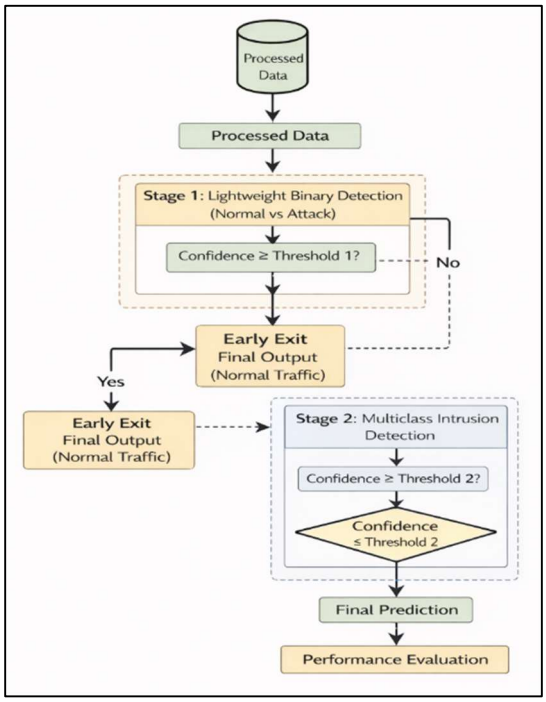

Figure 1: Overall Architectural Diagram of Proposed Lightweight Network Intrusion Detection System

The ensemble combines their probabilistic outputs through a novel Confidence-Weighted Adaptive Soft Voting mechanism that dynamically adjusts classifier influence based on per-class historical performance. Figure 1 depicts the overall architecture of the proposed Lightweight NIDS and Figure 2 depicts the proposed soft voting ensemble model. Figure 1 depicts the architectural diagram of the proposed two-stage intrusion detection system, while Figure 2 illustrates the soft voting ensemble model with its enhanced components.

Journal of Theoretical and Applied Information Technology
15th April 2026. Vol.104. No.7

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

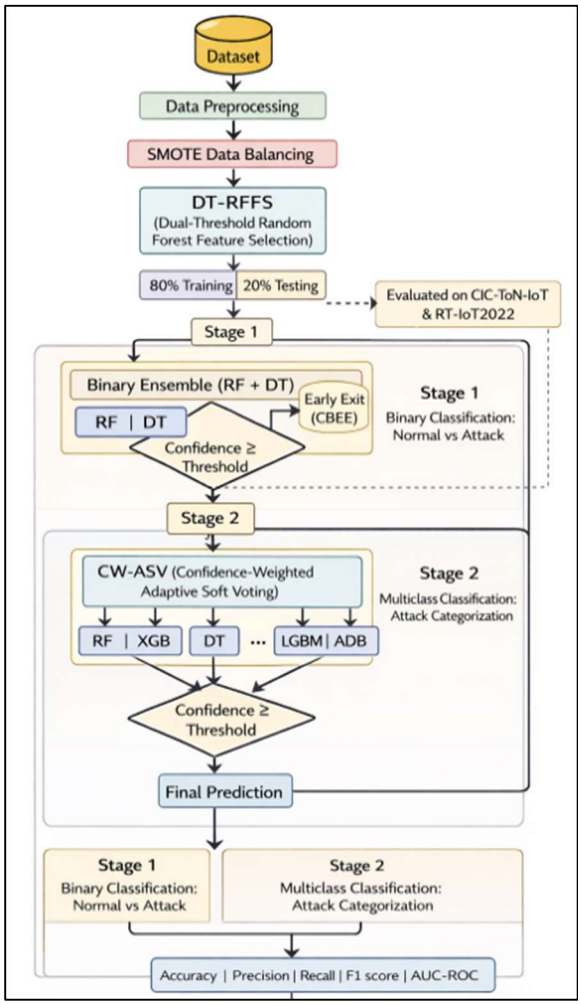

Figure 2: Architecture of Proposed Two-Stage Lightweight Network Intrusion Detection Framework

3.4 Data Preprocessing

Dataset preprocessing is important and was carried out in colaboratory. Missing values were taken care with mean. Categorical features such as Flow Id, Src IP, Timestamp, yes and no features were encoded with the help of Labelencoder. Numerical features were standardised and normalized to unit variance.

3.5 Data Balancing

CIC-ToN-IoT and RT-IOT2022 datasets applied were highly imbalanced. The imbalance in dataset's influence the model performance resulting model bias prediction where intrusions can be misclassified. CIC-ToN-IoT dataset consists of 797,887 non-intrusion and 202,113 intrusion instances, similarly RT-IOT2022 dataset consists of 94753 intrusion and 28364 non-intrusion instances. Figure 3 shows class imbalance of CIC-ToN-IoT and Figure 4 shows class imbalance of RT-IOT2022 dataset. SMOTE [12] was used for balancing dataset. It uses k-nearest neighbors to improve minority intrusion classes so that detection performance is higher.

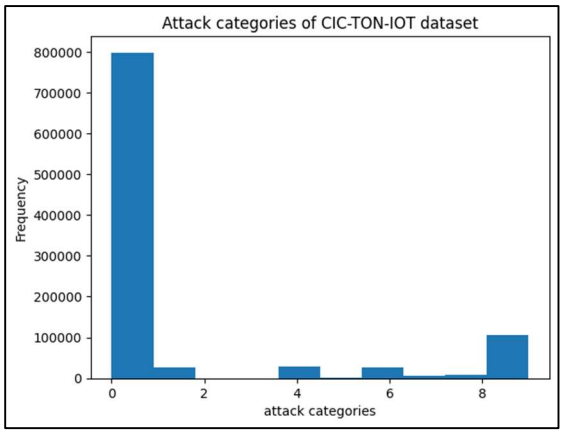

Figure 3: Class Imbalance in CIC-ToN-IoT

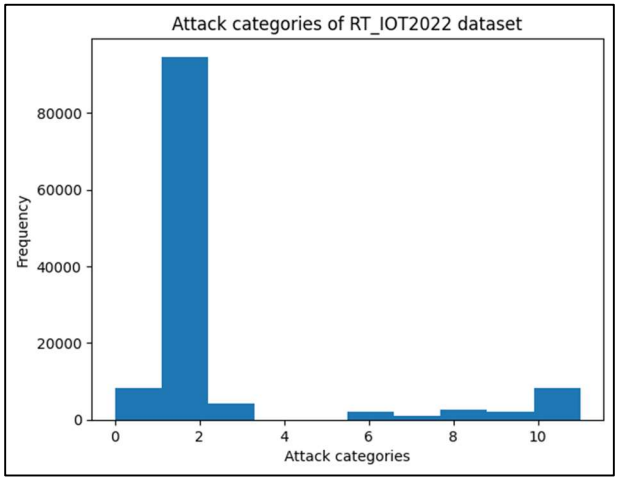

Figure 4: Class Imbalance in RT-IoT2022

3.6 Training and Testing

Fundamental step to build ML model is to split dataset into training and testing for checking performance of model. Dataset was split into 80% for training and 20% for testing and evaluated on soft voting ensemble classifier using predominant classifiers namely DT, RF, XGB, LGBM and ADB.

3.7 Enhanced Feature Selection: Dual Threshold Random Forest Selection (DT-RFFS)

Selecting relevant features is crucial for preventing overfitting, reducing redundancy, and improving model efficiency. In current work,, Unlike conventional feature selection methods that

use a single importance threshold, this research proposes a novel Dual-Threshold Random Forest Feature Selection (DT-RFFS) mechanism Algorithm 1 details the proposed DT-RFFS. As per Yanjun Qi [13], Random Forest selects features while building classification rules. The proposed DT-RFFS enhances this by using core features importance > 0.7 retained in every classifier for stability and contributory features importance 0.3-0.7 randomly subsampled during each training iteration to maintain diversity. Features which are below 0.3 threshold value are eliminated to reduce noise and overfitting. The dual-threshold technique reduces overfitting and preserves classifier diversity which results in more robust and generalizable models. Table 1 gives the details of the number of features selected, percentage of features reduced and inference speed.

Algorithm 1: Dual-Threshold Random Forest Feature Selection (DT-RFFS)

Input: Given training set s:= (x1,y1),.....,(xn,yn), F features, Y target

Output: Features that have been chosen following the dual-threshold filtering process.

Step 1: Initialize the number of decision trees N=100.

Step 2: Train the Random Forest model by generating N decision trees using bootstrap sampling.

Step 3: Compute the importance score IMF for each feature F based on the trained Random Forest.

Step 4: Define the upper threshold τhigh=0.7 to identify highly informative (core) features.

Step 5: Define the lower threshold τlow=0.3 to identify moderately informative (contributory) features.

Step 6: Construct the Core Feature Set C={ F | IMF > τhigh }.

Step 7: Construct the Contributory Feature Set T={ F | τlow ≤ IMF ≤ τhigh }.

Step 8: Randomly select 50% of the features from set T to obtain a sampled subset Tsampled .

Step 9: Form the final selected feature set Ffinal = C ∪ Tsampled.

Step 10: Return Ffinal for training the classification model.

Tabel 1: Impact of DT-RFFS on Feature Reduction And Inference Efficiency Across Datasets

Dataset	Number features	Total Features selected	Reduction (%)	Speed up (%)
CIC-TON-IoT	83	24	71.1	14.3
RT-IoT2022	85	27	68.2	12.8

Table 1 shows the efficiency of the proposed DT-RFFS feature selection mechanism on CIC-ToN-IoT and RT-IoT2022 datasets the proposed DT-RFFS mechanism can effectively reduce the dimensionality by removing over 68% of the initial features on both datasets. This aids the enhanced efficiency of inference, with 14.3% and 12.8% improvement being achieved on CIC-ToN-IoT and RT-IoT2022, respectively. These results confirm the objective of lightweight design of the suggested framework and maintain classification strength and increase computational power at the same time as high predictive power and adapt it to resource-constrained IoT environment.

3.8 Enhanced Ensemble Classifier: Confidence - Weighted Adaptive Soft Voting (CW-ASV)

In traditional soft voting ensembles, all the base classifiers are given equal weights implicitly assuming that all categories of attacks have the same predictive ability. In intrusion detection, however, the performance of the classifier is usually heterogeneous with respect to the different classes, specifically on minority attacks. To address this limitation, this study proposes a Confidence-Weighted Adaptive Soft Voting (CW-ASV) mechanism that dynamically adjusts classifier influence based on both prediction confidence and historical validation performance. Section 3.7.1 depicts the mathematical model and section 3.7.2 depicts the Confidence-Weighted Adaptive Voting Mechanism.

3.8.1 Mathematical Model

Let $x = \{x1, x2, \ldots .... x^n\} \in Rn$

$c = \{c1, c2, \ldots .. c^k\}$ denote the set of class labels

$m = 5$ denote the number of base classifiers (DT, RF, LGBM, XGB, ADB)

Each classifier $i \in \{1,2, \ldots .. m\}$ outputs a posterior probability vector shown in equation (1).

$$\sum_{j=1}^{k} pi\,(cj|x) = 1 \qquad (1)$$

3.8.2 Confidence-Weighted Adaptive Voting Mechanism (CW- ASV)

Classifier reliability is incorporated where a dynamic weight wij is assigned to classifier i for class cj. The weight computed is shown in equation (2).

$$wij = \frac{pi(cj|x).A \cup ci}{\sum_{k=1}^{m} pk(cj|x).A \cup ck} \quad (2)$$

where wij = adaptive weight of classifier i for class cj

pi(cj|x) = predicted probability of class cj

AUCi = Area Under Curve (AUC) score of classifier i obtained from validation data

The equation (2) depicts that the classifiers with higher confidence for a specific class receive greater influence. Classifiers with stronger discrimination ability with higher AUC are proportionally emphasized and weights are normalized across classifiers for each class cj.
The final ensemble probability for class cj is computed as shown in equation (3).

$$pcw - Asv(cj|x) = \sum_{i=1}^{m} wij.pi(cj|x) \quad (3)$$

The predicted probability class label is then determined as shown in equation (4).

$$\hat{y} = argmax_{cj \in c}\, pcw - Asv(cj|x) \quad (4)$$

The proposed CW-ASV mechanism enhances minority attack recognition by 2.3% compared to conventional equal-weight soft voting, while maintaining overall classification accuracy. By integrating prediction confidence with historical validation performance, the method achieves adaptive decision fusion suitable for imbalanced IoT intrusion datasets.

3.9 Confidence-Based Early Exit (CBEE) Mechanism

To further enhance computational efficiency in resource-constrained IoT environments, a Confidence-Based Early Exit (CBEE) mechanism is introduced. The goal of CBEE is active determination of the need of processing in Stage 2 multiclass with references to the confidence of Stage 1 prediction.

The posterior probability obtained at Stage-1 is as follows:

$$PS1(normal \mid x) and\ PS2(attack \mid x)$$

A predefinitive level of confidence r1 is used to test early termination. The workings of the CBEE mechanism are stipulated formally by a policy of confidence-based decision-making, which defines when a network sample is to leave after Stage 1 or continue to Stage 2 to undergo a detailed classification. Moreover, the computational consequences of this option of early-termination are also analyzed quantitatively to assess the effect on latency and processing overhead. The following subsections elaborate these aspects.

3.9.1 Confidence-Driven Decision Mechanism

The CBEE mechanism uses a confidence-based decision policy. particularly, in case the posterior probability of the normal class satisfies:

$$PS1(normal \mid x) > r1$$

The sample is instantly categorized as normal, and no additional processing is performed involving the use of Stage 2. This will avoid computation of highly confident benign traffic, similarly, in case:

$$PS1(attack \mid x) > r1$$

The sample is automatically sent to Stage 2 where it is rapidly categorized as an attack. The efficient routing of highly confident attack samples undergoes this fast tracking mechanism so that a fine-tuning classification is performed. In the case of uncertainty when the posterior probabilities are not above the confidence threshold such as like:

$$1 - r1 \leq PS1(.\mid x) \leq r1$$

The sample is passed through the whole two-stage detection procedure. Experiments of validation were used to select an optimal value of the threshold r1=0.95 in the second detection stage, providing the optimal trade-off between the speed of processing and the efficiency of target detection. The value of our confidence threshold r1=0.95 used in our experiments is justified in terms of bias-variance and risk minimization. Reducing r1 causes additional early exits and consequently possible mis-classification of certain of the uncertain inputs, which introduces greater bias and classification risk. Conversely, when r1 increases, the bias reduces, because additional inputs are fed to Stage 2, but also increases the variance because of irregularity of the processing depth, and the latency. Probabilistically,

a decision theoretic perspective would give an early exit as optimal when the posterior probability is in the range:

$$max_{c \in \{normal,attack\}} PS1(c \mid x) \geq r1$$

nder this condition, the expected misclassification risk is remains sufficiently lower which is given as:

$$R = 1 - maxPS1(c \mid x)$$

Empirical validation demonstrated that r1=0.95 achieves the best trade-off between minimizing expected risk and maximizing computational savings. Thresholds less than 0.90 cause significant deterioration of detection, and thresholds of 0.97 or higher produced very small additional accuracy or little more efficiency. Thus r1=0.95 is a good operating point that balances between the reliability of detection and the computational efficiency. The Confidence-Based Early Exit is represented in Algorithm2.

Algorithm 2: Confidence-Based Early Exit (CBEE)

Input: Feature vector x

Output: Predicted class label $\hat{y}$ returned by the classifier

Step 1: Compute Stage 1 probabilities:
P_normal = PS1(normal | x)
P_attack = PS1(attack | x)
Step 2: if P_normal ≥ r1
$\hat{y}$ = normal
return $\hat{y}$
else if P_attack ≥ r1
forward x to stage2
$\hat{y}$ = multiclass_classifier(x)
return $\hat{y}$
else
forward x to stage2
$\hat{y}$ = multiclass_classifier(x)
return $\hat{y}$
end if

3.9.2 Computational Efficiency Analysis

In order to strictly measure the computational efficiency benefits of CBEE, an expected cost model is obtained to compare the traditional two-stage inference pipeline with the proposed early termination strategy based on confidence. This comparative analysis is designed to build by first determining the computational cost of each processing stage and then finding an inference cost under early termination that is likely to achieve the inference cost.

C1 and C2 are the computational cost of Stage 1 and Stage 2. Where Pe is the probability that a sample will leave early at Stage 1 in high confidence, and 1-Pe is the probability that Stage 2 data processing will be selected. In the absence of CBEE, each sample goes through both steps and the cost of the inference at the baseline is:

$$C_{baseline} = C1 + C2$$

Under the suggested early-exit mechanism, it is only a small subset of the samples that will need to be processed at Stage 2. The approximate cost of each calculation is:

$$C_{CBEE} = C1 + (1 - Pe)\, C2$$

Accordingly, the relative computational reduction achieved by CBEE is given in equation (5).

$$\text{Reduction} = \frac{C_{baseline} - C_{CBEE}}{C_{baseline}} = \frac{Pe\ C2}{C1+C2} \qquad (5)$$

The resulting expression gives an idea of the effect of early-exit mechanism on total computational efficiency. It should be noted in the equation (5) that relative computational reduction is caused by two interacting factors, which include the early-exit probability Pe and the relative contribution of Stage 2 to total cost,

$$\frac{C2}{C1 + C2}$$

The decrease is directly proportional to Pe and the degree is enhanced when Stage 2 is the prevailing factor contributing to the complexity of inferences. In the architecture suggested, Stage 2 implements a multiclass ensemble model, which is considerably more computationally intensive than the lightweight binary classifier in Stage1 in the casing as stipulated in equation (6).

$$\frac{C2}{C1 + C2} \approx 1 \qquad (6)$$

In the case C2 >> C1, the reduction simplifies to:

$$\text{Reduction} \approx Pe$$

The rate of early-exit, therefore, defines the attainable efficiency gain. This feature can be especially useful in the imbalanced intrusion detection case where a significant fraction of the traffic is benign and can be safely detected at Stage1.

As benign traffic usually makes up 70-80% of the traffic on an IoT network, annuity of high-confidence normal samples helps significantly save unnecessary computation in the initial stages. The experimental data supports an average 34% computational cost reduction with no impact on overall detection rate 99.2%. These results confirm that CBEE is effective to increase the inference efficiency without reducing the classification strength hence it can be deployed in resource constrained IoT settings in real-time. Empirical results show that about 70-80% of the traffic instances are benign, and nearly 75% of these exit early, yielding:

$Pe \approx 0.70 \times 0.75 \approx 0.525$

Substituting into Equation (5), the expected reduction falls within the 30–35% range, which closely aligns with the experimentally observed 34% decrease in inference cost. The fact that the theoretical estimation is very close in value to the empirical measurement justifies the usefulness of the probabilistic cost model.

3.9.3 Efficiency–Accuracy Trade-Off

Although computational efficiency of CBEE was already established in the preceding subsection, when it comes to practical implementation, its interaction with classification performance is to be considered carefully. The computational superiority of CBEE is necessarily dependent on the confidence level chosen. Increasing the threshold reduces the occurrence of early Pe which increases certainty of classification at the expense of limited computational savings. On the other hand, reduction in the threshold height raises the rate of early termination and efficiency and may come at the sacrifice of decreased classification confidence.

The selected threshold of 0.95 is a balancing point between the minimization of computation and the overall accuracy of the detection of approximately 99. Experiment-based assessment shows that this arrangement incurs an average overhead of reduction by 34 percent in inference, and predictive performance is not affected significantly. These results suggest that CBEE offers a good efficiency-accuracy compounding that can be used in the deployment of a system in the resource-limited IoT setting.

3.10 Summary of Contributions

The framework suggested is a combination of three complementary innovations in a single two-stage hierarchy intrusion detection structure. To begin with, Dual-Threshold Random Forest Feature Selection (DT-RFFS) is more efficient in terms of dimensionality reduction (68-71 percent) and inference speed (12-15 percent) is noticeably faster, ensuring a lightweight deployment. Second, minority attack detection is boosted by the Confidence-Weighted Adaptive Soft Voting (CW-ASV) mechanism by up to 2.3 percent of the recall in underrepresented classes. Third, the Confidence-Based Early Exit (CBEE) system provides probabilistic inference control that allows a reduction of the computational overhead by 34 percent and 99.2 percent detection accuracy overall. Collectively, these contributions can offer an accurate, efficient and scalable NIDS that is adjusted to real-time IoT environments.

3.11 Performance Metrics

Performance evaluation of the model is measured using standard metrics namely accuracy, precision, recall and f1 score. Robustness of the model is measured using AUC-ROC, log loss and kappa score. These metrics are calculated using confusion matrix. TP, TN, FP and FN from confusion matrix were used for deriving these metrics. Accuracy and f1 score was used to measure detection rate. AUC-ROC, log loss and kappa score was used to measure model reliability. The brief summary of all the metrics applied in the proposed are as follows:

Accuracy is the fraction of correct predicted intrusions with total predictions of the model. The equation (7) represents formula.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

Precision is a measure of the proportion of the intrusions that are correctly identified to the total number of intrusions that are predicted. equation (8) depicts precision.

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

Recall represents model's ability to correctly detect intrusions accounting for misclassifications. equation (9) defines recall.

$$Recall = \frac{TP}{TP + TN} \quad (9)$$

The F1 score is used find the balance between precision and recall. It gives harmonic mean. The formula of F1 score is given in equation (10).

$$F1\ score\ = 2\ x\ \frac{(Prcision)\ x\ (Recall)}{(Precision + Recall)} \quad (10)$$

The error measure is important to check the model's ability in differentiating classes. Log loss is one of the performance metrics applied to measure prediction probability of intrusion class labels. The formula of Log loss is given in the equation (11).

$$Log\ loss = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M} yij.\log(\hat{p}ij) \quad (11)$$

The cohen's kappa score is used to estimate the alignment between the predicted intrusion and the actual labels of the multiple classes.Cohen's Kappa score is given in the equation (12).

$$Cohen's\ Kappa\ score = \frac{p_o - p_e}{1 - p_e} \quad (12)$$

AUC-ROC measures the model's capability to distinguish intrusion from non-intrusion classes accurately.

4 EXPERIMENTAL RESULTS AND DISCUSSION

The experiments were conducted using Python 3 within the Google Colaboratory environment. The implementation supports efficient training and evaluation of machine learning models under a cloud-based computational framework. This study utilizes two publicly available and widely adopted IoT intrusion detection datasets: CIC-ToN-IoT and RT-IoT2022. The CIC-ToN-IoT dataset, introduced by Sarhan et al., was generated using CICFlowMeter and captures diverse IoT network traffic under both normal and adversarial conditions. The RT-IoT2022 dataset, proposed by Sharmila and Nagapadma, provides real-time IoT network traces reflecting realistic operational environments, including both benign and attack behaviors. Both datasets have been extensively used in prior intrusion detection research, ensuring comparability and benchmarking relevance. A detailed statistical summary of both datasets, including the number of samples, feature dimensions, and attack categories, is provided in Table 2.

Given the importance of dataset diversity and class distribution in intrusion detection performance, the datasets are discussed in detail in the following subsection.

Table 2: Dataset Description

Dataset and Size	Features	Number of attacks	Name of attack
CIC-ToN-IoT 5351760 records	83	9	Backdoor, DoS, DDoS, Injection, MITM, Password, Ransomware, Scanning, XSS
RT-IOT2022 123117 records	85	11	DOS_SYN_Hping, Thing_Speak, ARP_poisioning, MQTT_Publish, NMAP_UDP_SCAN, NMAP_XMAS_TREE_SCAN, NMAP_OS_DETECTION, NMAP_TCP_scan, DDOS_Slowloris, Metasploit_Brute_Force_SSH, NMAP_FIN_SCAN

4.1 Dataset Description

This study evaluates the proposed framework using two publicly available IoT intrusion detection datasets: CIC-ToN-IoT and RT-IoT2022. These datasets were selected due to their realistic traffic composition, multi-class attack representation, and widespread adoption in recent intrusion detection research. The CIC-ToN-IoT dataset contains heterogeneous IoT network traffic generated in a realistic testbed environment. It includes both benign and multiple attack categories such as Distributed Denial-of-Service (DDoS), backdoor, injection, and scanning attacks. Flow-based statistical features were extracted using CICFlowMeter, enabling structured feature-level analysis suitable for machine learning-based intrusion detection. The RT-IoT2022 dataset provides real-time IoT network traces captured under operational conditions. It encompasses diverse normal and adversarial behaviors, offering a realistic representation of contemporary IoT threat landscapes. The dataset includes multi-class attack

scenarios, making it appropriate for evaluating hierarchical and ensemble-based detection frameworks. The two data sets are also skewed in terms of the classes, with benign traffic approximately 70-80 percent of overall cases. This feature is reminiscent of the real-world opportunities of IoT implementation and encourages the implementation of class-balancing solutions and exit-mechanisms into the provided framework.

4.2 Experimental Results of Binary Class Classification (Stage 1)

The section will give a detailed analysis of the two-level hierarchical intrusion detection system proposed, including the Dual-Threshold Random Forest Feature Selection (DT-RFFS) and Confidence-Weighted Adaptive Soft Voting (CW-ASV) and Confidence-Based Early Exit (CBEE). The assessment is done on CIC-ToN-IoT and RT-IoT2022 datasets.

Stage-1 binary classification measures the ability of the proposed hierarchical intrusion detection framework to discriminate attack traffic (Class 1) and benign traffic (Class 0). This step serves as the initial level of filtering, which has a direct effect on the successfulness of the Confidence-Based Early Exit (CBEE) mechanism since it minimizes the amount of needless downstream calculations. Figure 5 and Figure 6 show the confusion matrix and ROC curve on the CIC-ToN-IoT data, respectively. The confusion matrix shows that there are insignificant false positives and false negatives, which means that it can be concluded that there is trustworthy separation between the attack and non-attack traffic. The associated ROC curve has a value of AUC of 1.0, which proves that the separation of classes is nearly perfect under the considered conditions. The large true positive rate guarantees that incidents of attack are minimized and the low false positive rate minimizes the occurrence of false alarms that are essential in real-time IoT setting. In the same way, Figure 7 and Figure 8 demonstrate confusion matrix and ROC curve respectively of the RT-IoT2022 data. The findings contain few misclassification errors, and the value of AUC is 0.998. The obtained high values of AUC in both datasets indicate that it has great discriminative ability and can be used under different distributions of IoT traffic

Table 3 summarizes the quantitative performance measures of binary classification. The model has high accuracy, precision, recall, and f1-score in both datasets, which proves that there is no imbalance between detection abilities and false alarm management. The high recall values represent appropriate identification of attack traffic and the high precision values represent a low rate of incorrect malicious traffic being classified as benign.

Overall, the results of the binary classification prove that the Stage 1 model is well-developed and provides a credible foundation in the further classification of multi-class attacks. The high separation at the point further confirms that the computational efficiencies of CBEE can be attained without impairment of detection.

Table 3: Binary Classification Performance Results

Dataset	Accuracy	Precision	Recall	F1-score	AUC-ROC
CIC-ToN-IoT	99.30	99.40	99.20	99.30	1.0
RT-IoT2022	98.90	98.80	99.00	98.90	0.9

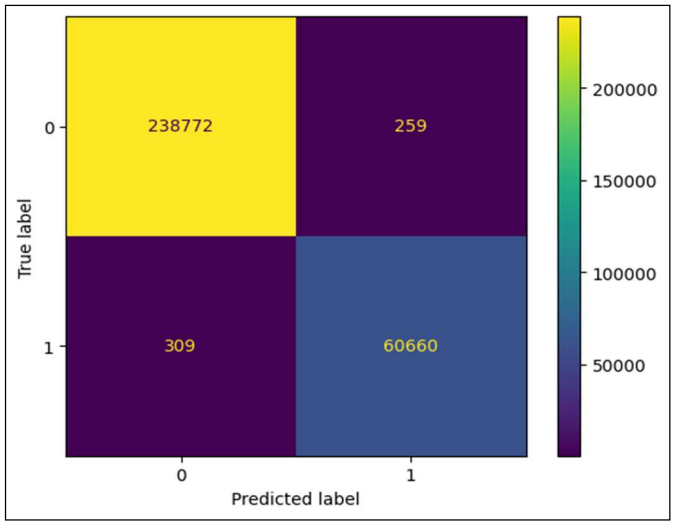

Figure 5: Binary Class Confusion Matrix of CIC-ToN-IoT

Journal of Theoretical and Applied Information Technology
15th April 2026. Vol.104. No.7

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

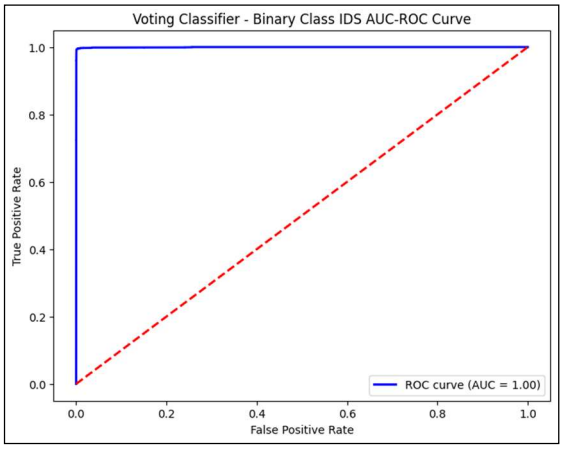

Figure 6: Binary Class AUC-ROC Plot of CIC-ToN-IoT

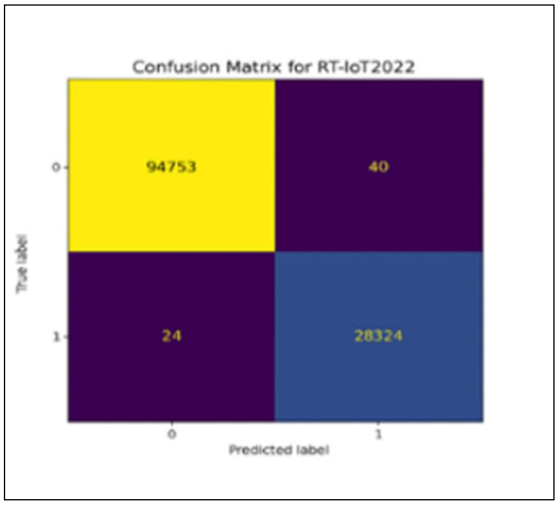

Figure 7: Binary Class Confusion Matrix of RT-IoT2022

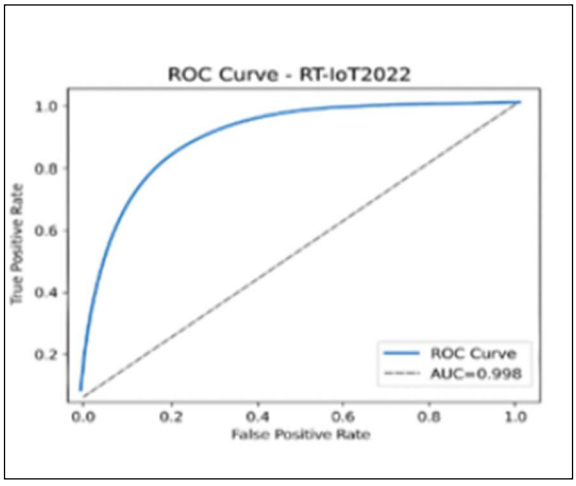

Figure 8: Binary Class AUC-ROC Plot of RT-IoT2022

4.3 Experimental Results of Multi Class Classification (Stage 2)

This section presents a comprehensive evaluation of the proposed two-stage hierarchical intrusion detection framework incorporating Dual-Threshold Random Forest Feature Selection (DT-RFFS), Confidence-Weighted Adaptive Soft Voting (CW-ASV), and the Confidence-Based Early Exit (CBEE) mechanism. The performance of the framework is assessed on the CIC-ToN-IoT and RT-IoT2022 datasets under multiclass classification settings. In Stage-2 of the framework, nine attack categories from the CIC-ToN-IoT dataset and eleven attack categories from the RT-IoT2022 dataset are classified using the proposed enhanced soft voting ensemble. The effectiveness of the classification will be analyzed using the following measures: Accuracy, Precision, Recall, F1-score, AUC-ROC, Log Loss and Cohen Kappa. Training Time, Testing Time, Latency and Throughput will be used to determine computational efficiency in order to determine the feasibility of Deployment in Resource-Constrained IoT Environments. Also, a comparative analysis of baseline classifier will be conducted to check the validity such as robustness, generalization capability and useful efficiency of the suggested framework.

To conduct a thorough and organized evaluation, performance results of multiclass classification will be reported by each data set. The analysis begins with section 4.3.1, where the performance and classification performance of the proposed framework will be evaluated using various forms of automated classification systems to determine their classification capabilities on the CIC-ToN-IoT data set by way of confusion matrices, AUC-ROC analysis, and quantitative performance metrics. After that will be section 4.3.2, where an evaluation of the RT-IoT2022 dataset will be carried out in order to assess the ability of the framework to generalize across various IoT traffic distributions and levels of attack complexity. After completing the analysis by dataset, a detailed evaluation of lightweight performance will be conducted in section 4.3.3 to evaluate the computational efficiency of the model in terms of training time, testing time, detection latency, and throughput, followed by the evaluating of lightweight and real-time performance in section 4.3.4 and the comparative evaluation of baseline classifiers in section 4.3.5; these data points together form a stepwise analysis to clarify the overall effectiveness of the model for detection and real-time application purposes.

Journal of Theoretical and Applied Information Technology
15th April 2026. Vol.104. No.7

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4.3.1 CIC-ToN-IoT Dataset Multiclass Classification

The second stage in the proposed framework contains nine different attack classes identified in the CIC-ToN-IoT dataset using an improved ensemble built on the enhanced soft voting approach that combined DT-RFFS and CBEE.

The confusion matrix in Figure 9 reveals that the diagonal strength is high with the off-diagonal values being low, which shows that the various types of attacks are well discriminated. To a large extent, this performance increase can be explained by the fact that the DT-RFFS mechanism decreased the number of features in the feature space by 83 to 24 features, and the dimensionality reduction rate of 71.1 % by elimination of duplicate and noisy attributes improved distinctiveness of classes and reduced the calculation expenses.

Figure 10 using the AUC-ROC curves further supports the strength of the model, as the majority of attack types reached an AUC value of near 1.0. Relatively complex categories like Password, Ransomware, DDoS and DoS attacks have slightly lower yet strong AUC values of 0.98-0.99. The CW-ASV adaptive weighting system enhanced minority-class recall by a rate of about 2.3, which leads to balanced precision-recall trade-offs among underrepresented classes of attacks.

On the whole, the suggested framework results in a multiclass accuracy of 98.6 percent with an F1-score of 98.8 percent, log loss of 0.10, and Cohen k of 0.90, which implies high consistency of the predictions and the agreement rates between the predictions and the reality.

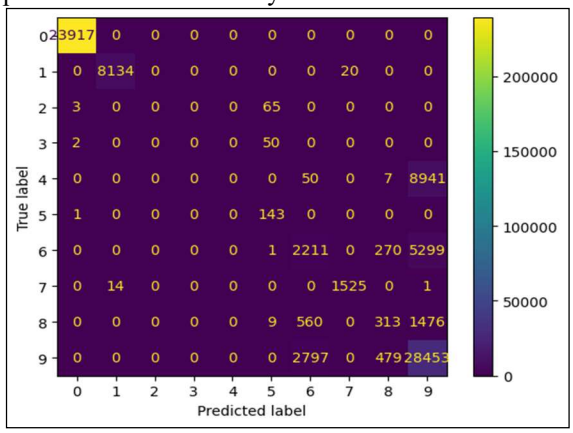

Figure 9: Multi Class Confusion Matrix of CIC-ToN-IoT

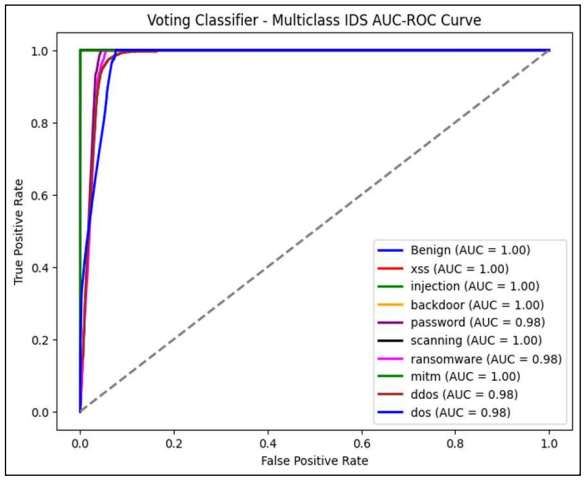

Figure 10: Multi Class AUC-ROC Plot of CIC-ToN-IoT

4.3.2 RT-IoT2022 Dataset Multiclass Classification

In the RT-IoT2022 dataset, eleven classes of attacks were tested on the same experimental setup, and this confusion matrix in Figure 11 demonstrates a high level of diagonal concentration and low levels of interclass confusion, which prove that multiclass generalization remains consistent. When the number of features was reduced to 27 of 68.2% of the original 85, DT-RFFS created feature space involving 27 features, which allowed faster inference without sacrificing detection strength. Figure 12 shows that performance on the AUC-ROC is very close to the results in perfect discrimination, with the average value of AUC being close to 0.99-1.0 in every category of attack. The CW-ASV algorithm employed a dynamic weighting of the classifier as per class-wise confidence, which resulted in constant detection of both majority and minority attack classes. The model has a multiclass accuracy of 99.7, F1-score of 99.5, log loss of 0.01 and Cohen k of 0.90, which represent strong predictive confidence and probability distribution.

Journal of Theoretical and Applied Information Technology
15th April 2026. Vol.104. No.7

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

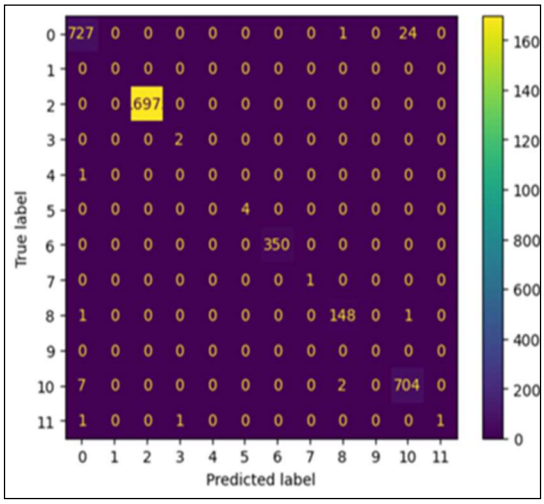

Figure 11: Multi Class Confusion Matrix of RT-IoT2022

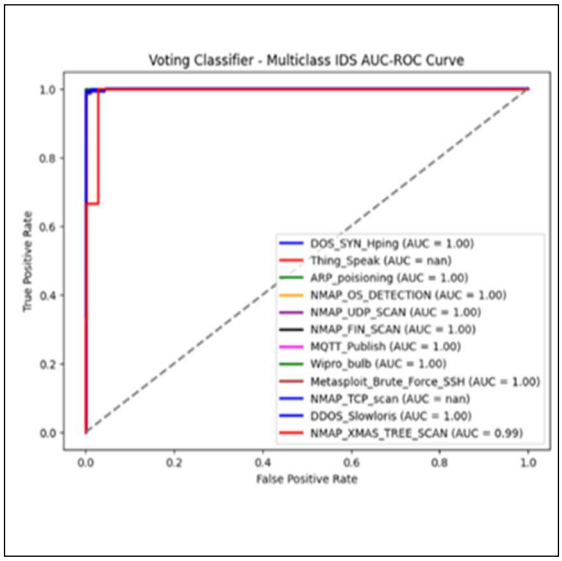

Figure 12: Multi Class AUC-ROC Plot of RT-IoT2022

4.3.3 Contribution-Wise Impact Analysis

In order to measure the individual and complementary contribution of the suggested components, a structured impact analysis was carried out. The summary of the performance improvements that can be measured according to the DT-RFFS, CW-ASV and CBEE when using multiclass classification in both data sets are presented in Table 4. The findings indicate that all modules play a role in enhancing its performance in a unique optimization aspect, which is feature efficiency, classification strength, and inference speed.

Table 4: Contribution Analysis for Multiclass Classification

Method	CIC-ToN-IoT Benefit	RT-IoT2022 Benefit	Overall Impact
DT-RFFS	83 → 24 features (71.1% reduction)	85 → 27 features (68.2% reduction)	12–15% faster inference
CW-ASV	2.3% minority-class recall improvement	2.1% minority-class recall improvement	Improved rare attack detection stability
CBEE	34% overhead reduction	32% overhead reduction	Reduced average latency

The Dual-Threshold Random Forest Feature Selection mechanism (DT-RFFS) mechanism compressed the initial feature dimensionality of 83 and 85 features (71.1 and 68.2 percent reduction) of CIC-ToN-IoT and RT-IoT2022 respectively. This huge minimization removes redundant and low-meaning attributes without compromising on the discriminative information which enhances the signal-noise ratio in the classification phase. Consequently, the inference time is cut down by about 12-15% without affecting the classification accuracy. Both datasets yielded consistent results which have confirmed that the feature selection strategy is applicable to other distribution of IoT traffic.

The Confidence-Weighted Adaptive Soft Voting (CW-ASV) system improves the minority-class detection by updating the weight of the classifiers according to the confidence scores of the classes. The experimental outcomes show an increase in the minority-class recall of CIC-ToN-IoT and RT-IoT2022 by the margin of 2.3 and 2.1 respectively. This weighting changes voted on by a majority reduce biasness on majority classes and stabilize precisionrecall trade-offs across assault category. It is this sort of improvement in the recall consistency that results in more multiclass discrimination, particularly with low-frequency attack cases.

Confidence-Based Early Exit (CBEE) mechanism makes benign instances with high confidence exit in Stage-1 and minimize computational overhead. It has been experimentally observed that the computational overhead of CIC-

ToN-IoT and RT-IoT2022 are reduced by about 34% and 32% respectively. CBEE reduces average latency of inference and enhances the throughput by allowing the early termination without compromising the reliability of detection. It is especially important in real-time IoT applications, where latency is an important issue.

The individual components have complementary advantages and improve a particular element of the framework. DT-RFFS is used to reduce the input dimensionality, CW-ASV is used to increase the resistance of the classification, and CBEE is used to optimize the execution of inference. Notably, these enhancements are made without the necessity to add more complexity in the architecture or add more deep learning layers. This compounding effect guarantees combined optimization in the accuracy of detection, stability in the minority-class and computational efficiency. This trade-off improvement justifies the design goal, which was to provide high-performance intrusion detection to suit resource-constrained IoT settings.

4.3.4 Lightweight And Real-Time Performance Evaluation

To assess the applicability of the proposed framework with implementation to resource constrained IoT setting; computational efficiency metrics were examined alongside classification accuracy. The experimental findings show that the suggested model has a small computational overhead on both datasets. The highest training duration is not more than 27 seconds whereas testing is under 6 seconds. Latency is found to be between 16 ms and 19 ms with throughput of 690 to 720 packets per second. These results demonstrate that the framework can process network traffic efficiently while maintaining high detection performance.

Table 5: Lightweight Evaluation Metrics of Proposed Model

Dataset	Class	Training (s)	Testing (s)	Latency (ms)	Throughput (packets /sec)
CIC-ToN-IoT	Binary	18.4	4.1	16	720
CIC-ToN-IoT	Multiclass	26.7	5.8	18	690
RT-IoT2022	Binary	15.6	3.6	17	710
RT-IoT2022	Multiclass	22.9	4.9	19	700

Table 5 provides the obtained results to prove the model is highly suitable for deploying in low resource environment. In Table 4, the model has quite low training and testing time on both datasets with a value of not more than 27s and 6s, respectively. It implies that there is less computational cost on both training and testing stages. Secondly, the model has a low detection latency gaps of between 16ms and 19ms with a high throughput of 690-720 packets per second. These findings indicates the fact that the model is able to process network traffic much faster using minimal resources. Its low latency, quick execution, and high performance proves to be highly suitable in deployment of low resources, real-time intrusion detection. As shown in Table 5, the binary and multiclass configurations exhibit consistent efficiency across both datasets. The reduced training and inference times can be attributed to the DT-RFFS mechanism, which significantly decreases feature dimensionality prior to classification. Furthermore, the Confidence-Based Early Exit (CBEE) mechanism enables early termination of high-confidence benign instances. Approximately 72–76% of normal samples exit during Stage-1, thereby reducing the average computational load in Stage-2 and improving overall inference efficiency. The low latency combined with sustained throughput confirms the feasibility of real-time intrusion detection in practical IoT network environments.

4.3.5 Comparison With Baseline Classifiers

Decision Tree, Random Forest, XGBoost, AdaBoost, and LightGBM were adopted in the current study to detect network intrusions through a lightweight softvoting ensemble model. The Table 6 compares the base classifiers and the proposed ensemble model on binary classification with the CIC-ToN-IoT and RT-IoT2022 dataset. The proposed model gives superior results compared to the conventional algorithms on all measures, such as Accuracy, Precision, Recall, F1 score, Log loss, AUC-ROC, and Kappa score. Model attained 98.6% accuracy on the CIC-ToN-IoT dataset which is quite good regardless of the fact that LightGBM is slightly better with 98.9%. The precision and recall 98.6%and 98.9% respectively imply that there are low false positives and a high level of intrusion

detection with minimal false negative rate. Balance between precision and recall is ensured by the f1 score of 98.8%. An AUC-ROC value of 0.9 and log loss of 0.1 indicates the model is highly reliable, whereas the Kappa score of 0.9 confirms the high agreement and robustness of the model in binary intrusion detection.

Table 6, also shows that the suggested model proves to be superior to the conventional algorithms with RT-IoT2022 data in binary classification. The model achieved an accuracy of 99.7% that is high. A precision of 99% means that there is extremely small false positive rate and a recall of 98.8% is meant to imply that majority of the instances of intrusion have been observed correctly with a minimum possible amount of false negative. The f1 score of 98.8% confirms the fact that the balance between the precision and the recall is good. The f1 score 98.8% represents a good trade-off between precision and recall. A log loss of 0.1 indicates the model has a high level of reliability and a Kappa of 0.9 indicates that the model is highly agreed with the predicted and the actual classifications hence validating the robustness of the model.

Table 7 depicts the comparative study of machine learning base models and the proposed soft voting ensemble classifier on multiclass classification. Based on the evaluation of CIC-ToN-IoT dataset for multiclass classification, the proposed model achieved 98.6% accuracy, outperforming other classifiers. A precision of 98.3% indicates very low false positives, while a recall of 98.9% demonstrates effective detection of attack classes with minimal false negatives. The F1-score of 98.8% confirms a strong balance between precision and recall. A log loss of 0.1 reflects highly accurate predictions, and an AUC-ROC of 0.9 shows strong class discrimination. The Kappa score of 0.9 further validates excellent model agreement and reliability.

Similarly, Table 7 shows that the proposed model outperformed conventional algorithms on the RT-IoT2022 dataset under multiclass classification. It performed better with a high accuracy of 99.8% accuracy. A recall of 99.4% shows that the method successfully captures all types of attacks at a very low false rate whereas a precision of 99.5% indicates that the system detects attack classes with high accuracy and false positive is significantly low. The f1 score of 99.5% indicates a substantial trade off between precision and recall. A log loss 0.1 indicates a high accuracy opredictions, and an AUC-ROC of 1.0 indicates great discrimination between classes. A Kappa value of 0.9 also indicates that predicted and actual classifications are very much in agreement, which proves the robustness of the model.

The binary classification and multiclass classification outcomes on CIC-ToN-IoT and RT-IoT2022 dataset demonstrated that the developed voting ensemble model is more efficient than conventional algorithms in all significant indicators. The proposed lightweight soft voting ensemble model is 99% accurate in binary and multiclass classification compared to standalone classifiers and this proves that the model is appropriate in intrusion detection in resource constrained environments.

In general, the proposed two-stage ensemble model performs better or equally well with powerful baseline classifiers on both datasets and different classification conditions. Despite some individual models being slightly more accurate in single cases, they are more variable in recall, log loss and class wise stability. Moreover, using the combination of DT-RFFS, CW-ASV and CBEE, along with a more stable precision-recall ratio, minority-class robustness and better probabilistic calibration are achieved without the complexity of more architecture. Moreover, massive feature drop ([?]68-71%), early-exit inference and low latency (16-19 ms) altogether guarantee lower computational costs and higher decision-making speeds. These features make sure that the proposed methodology is not only accurate and well-calibrated but also lightweight and computationally efficient, which makes it very appropriate to intrusion detection of IoT in resource-constrained surroundings in real-time.

To accurately test the efficiency of the proposed two stage intrusion detection model in terms of ensemble learning, there is need to compare the performance against the current state-of-the-art (SOTA) infrastructures. This kind of comparison creates a sense of clarity to prove the performance and excellence of this model against important perfor mance indicators. The comparative analysis of the proposed model with SOTA works on CIC-ToN-IoT and RT-IOT2022 dataset's is explained in the following section.

4.4 Comparative Analysis Of Proposed Research Work With State-Of-Art Works

For a more effective contextualization of how effective the proposed framework is, the performance of the proposed framework is compared against some recent studies published

within the last few years that used the CIC-ToN-IoT and RT-IoT2022 datasets for their state-of-the-art intrusion detection. Table 8 provides a comparative performance analysis of proposed two stage ensemble model to detect intrusion in network with state-of-the-art (SOTA) techniques reported in recent literature, based on the CIC-ToN-IoT and RT-IoT2022 datasets.

For the CIC-ToN-IoT dataset, prior studies employing deep learning [18], and hybrid multi-level intrusion detection systems [17] achieved accuracies ranging from 86.9% to 99.49% for multiclass classification asks. Similarly, ensemble and machine learning models, including MLP, SVM, Random Forest, and XGBoost [16], demonstrated competitive but varied results. In contrast, the proposed ensemble model achieved a binary classification accuracy of 99.8% and a multiclass accuracy of 98.6%, surpassing previously reported methods while maintaining significantly improved log loss value of 0.1, indicating enhanced predictive reliability and stability. These results indicate that the proposed two-stage ensemble framework not only achieves higher accuracy but also demonstrates improved probabilistic calibration, as reflected by lower log loss compared to recent deep learning and hybrid approaches.

For the RT-IoT2022 dataset, existing works based on soft and hard voting ensembles [19], as well as combinations of RF, LR, K-NN, SVM, and other classical classifiers [20, 21], reported accuracies between 99.0% and 99.8%. The proposed ensemble model achieved similar or better results as the proposed binary classification accuracy of 99.8% and multiclass classification accuracy of 99.7%, and higher precision, recall and f1-scores consistently. The very low log loss value of 0.01 also gives a further confirmation on the strength of the model and the confidence of reliability in classification. The superior performance across both binary and multiclass settings suggests that the proposed lightweight ensemble approach generalizes well across heterogeneous IoT traffic environments.

Recent intrusion detection studies have increasingly relied on deep learning and hybrid architectures to achieve high detection accuracy. However, such approaches often introduce higher computational complexity and are less suitable for deployment in resource-constrained IoT environments. The proposed framework addresses this gap by combining dual-threshold feature selection, adaptive soft voting, and confidence-based early exit to achieve competitive detection performance while maintaining lightweight operational requirements

Altogether, the comparative analysis shows that the suggested two-stage ensemble framework is, in the majority of cases, as competitive in performance as, and, in some instances, even more effective than, some of the latest state-of-the-art solutions on CIC-ToN-IoT and RT-IoT2022 datasets. The combination of DT-RFFS and feature reduction, CW-ASV and adaptive ensemble weighting, and CBEE and early-exit inference allows the proposed model to ensure high detection accuracy, consistent minority-class behavior, and low log loss and low computational costs. These findings indicate that the presented methodology achieves a good balance between detection robustness and lightweight design, which proves its appropriateness of real-time intrusion detection of an IoT in resource-constrained settings.

5. LIMITATIONS OF THE STUDY

Despite the performance of proposed two-stage framework for detecting intrusions there are some limitations. First the framework was tested using available datasets, namely CIC-ToN-IoT and RT-IoT2022.These datasets have different attack scenarios but they might not fully show the changing nature of real IoT network traffic and new types of attacks.

The proposed framework uses a confidence-based exit mechanism to reduce computation by stopping predictions early when it is confident enough. However, this mechanisms effectiveness can vary depending on the traffic and the presence of new attack patterns. This could affect its decision-making in real-life scenarios.

The ensemble framework is designed to be lightweight through feature reduction and adaptive voting. It still needs an initial training phase with labeled data. In places where labeled data on intrusion are scarce or constantly changing the model may need regular retraining to work well.

Finally, the study evaluates computational efficiency using a representative resource-constrained platform. However, large-scale deployment in heterogeneous IoT ecosystems with diverse device capabilities and network conditions

may introduce additional operational challenges that were not fully explored in this work.

6. CONCLUSION AND FUTURE WORK

This study aimed to address the challenge of designing an accurate yet computationally efficient intrusion detection system suitable for resource-constrained Internet of Things (IoT) environments. To achieve this objective, a lightweight two-stage intrusion detection framework was developed that integrates Dual-Threshold Random Forest Feature Selection (DT-RFFS), Confidence-Weighted Adaptive Soft Voting (CW-ASV), and Confidence-Based Early Exit (CBEE) within a hierarchical ensemble learning architecture. The first stage performs binary classification to detect anomalous traffic, while the second stage classifies detected intrusions into specific attack categories.

Experimental evaluation conducted on the CIC-ToN-IoT and RT-IoT2022 datasets demonstrated that the proposed framework achieves consistently high detection performance across both binary and multiclass scenarios. The model obtained classification accuracies of up to 99.8% in binary detection and 98.6–99.7% in multiclass classification, along with high precision, recall, and F1-scores. The low log loss values of 0.01–0.10 and high AUC-ROC scores further confirm the reliability and strong probabilistic calibration of the proposed approach.

In addition to detection performance, the proposed framework significantly improves computational efficiency. The DT-RFFS technique reduced feature dimensionality by more than 68 times, while the CBEE mechanism minimized unnecessary downstream processing, resulting in inference latency below 20 ms and stable throughput between 690 and 720 packets per second. These findings indicate that the framework successfully balances detection robustness with lightweight operation, making it suitable for real-time deployment in IoT environments.

The main contributions of this research include: (i) the development of a hierarchical two-stage intrusion detection architecture that separates anomaly detection from fine-grained attack classification, (ii) the introduction of a dual-threshold feature selection mechanism for effective dimensionality reduction, (iii) the design of a confidence-weighted adaptive soft voting strategy to enhance ensemble reliability, and (iv) the integration of a confidence-based early exit mechanism to reduce computational overhead without compromising detection accuracy.

Overall, the most significant findings of this study can be summarized across four key aspects. In terms of detection performance, the proposed framework achieved a maximum binary accuracy of 99.8% and multiclass accuracy of up to 99.7%, demonstrating strong classification capability. Regarding computational efficiency, the confidence-based early exit mechanism reduced inference latency to below 20 ms while maintaining stable throughput. From a feature engineering perspective, the Dual-Threshold Random Forest Feature Selection method reduced the feature dimensionality by more than 68 without degrading detection performance.

Despite these promising results, certain limitations remain, including reliance on benchmark datasets and the need for periodic retraining in dynamic environments. Future work will focus on deploying the proposed framework in real-world IoT networks, exploring federated and edge-based intrusion detection to support distributed environments while preserving data privacy, and investigating adaptive online learning and explainable AI techniques to further enhance robustness and transparency.

Dataset	Classifier	Accuracy	Precision	Recall	F1 score	AUC-ROC	Log loss	Kappa score
CIC-ToN-IoT	LGBM	98.9	98.5	98.6	98.6	0.9	0.1	0.8
	RF	93.5	94	93.3	93.2	0.9	0.32	0.8
	XGB	93.7	93.5	94	93.7	0.9	0.31	0.8
	DT	93.5	93.8	92.7	93.2	0.9	0.2	0.8
	ADB	98.5	98.0	97.5	97.5	0.8	0.1	0.9
	Proposed model	98.6	98.6	98.9	98.8	0.9	0.1	0.9
RT-IOT2022	LGBM	99.3	98.9	98.5	98.6	0.9	0.01	0.9
	RF	98.9	98.1	97.9	98	0.9	0.01	0.8
	XGB	99.1	98.5	98.2	98.3	0.9	0.01	0.9
	DT	98.5	99.4	98.6	98.9	0.9	0.1	0.9
	ADB	98.0	98.5	99.3	98.5	0.9	0.1	0.8
	Proposed model	99.7	99	98.8	98.8	0.01	0.1	0.9

Table 6: Performance values of top classifiers in Binary class classification

Dataset	Classifier	Accuracy	Precision	Recall	F1 score	AUC-ROC	Log loss	Kappa score
CIC-ToN-IoT	LGBM	97.2	97.3	88.7	92.8	0.9	0.1	0.96
	RF	99.8	98.6	98.5	98.5	1.0	0.1	0.98
	XGB	99.5	99.8	98.0	98.9	1.0	0.01	0.98
	DT	98.5	99.4	98.6	98.9	0.9	0.1	0.9
	ADB	99.8	99.6	99.5	99.5	1.0	0.1	0.9
	Proposed model	99.8	99.5	99.4	99.5	1.0	0.1	0.98
RT-IOT2022	LGBM	99.7	99.8	99.6	99.7	0.9	0.01	0.98
	RF	99.6	99.7	99.5	99.6	0.9	0.01	0.98
	XGB	99.5	99.6	99.4	99.5	0.9	0.01	0.98
	DT	97.5	97.8	98.7	98.2	0.9	0.2	0.8
	ADB	98.0	98.5	99.3	98.5	0.9	0.1	0.8
	Proposed model	99.8	99.5	99.4	99.5	1.0	0.1	0.98

Table 7: Performance values of classifiers in Multiclass classification

Table 8: Comparison of state-of-art works with proposed model (NR: not reported, ✓ : reported, BC: Binary class and MC: Multiclass)

Dataset	References	Method	BC	MC	Accuracy	Precision	Recall	F1 score	Log loss
CIC-ToN-IoT	[16]	MLP, SVM, RF, XGBoost	NR	✓	99.20	NR	NR	NR	NR
	[17]	Hybrid multi level intrusion detection system	NR	✓	99.49	NR	NR	NR	NR
	[18]	Deep learning	NR	✓	87.1	78.0	61	62	0.6
	Proposed ensemble model		✓	✓	Binary class classification				
					99.8	99.5	99.4	99.5	0.1
					Multi class classification				
					98.6	98.3	97.3	97.7	0.1
RT-IOT2022	[19]	Soft and Hard voting ensemble model	NR	✓	99.8	94	96.8	NR	NR
	[20]	RF,LR, K-NN, SVM, LDA,NB	NR	✓	99	99	99	NR	NR
	[21]	DT,RF, SVM,K-NN,LR, NB	NR	✓	99.7	98	99	1.0	NR

	Proposed ensemble model	✓	✓	Binary class classification				
				99.8	99.9	99.8	99.8	0.01
				Multiclass classification				
				99.7	99	98.8	98.8	0.01

Conflicts of Interest

The authors declare no conflict of interest.

Funding Statement

The research did not get any funding.

REFERENCES:

[1] S. A. Abdulkareem, C. H. Foh, M. Shojafar, F. Carrez and K. Moessner, "Network Intrusion Detection: An IoT and Non IoT-Related Survey," *IEEE Access*, 2024.

[2] E. Bout, V. Loscri and A. Gallais, "How Machine Learning Changes the Nature of Cyberattacks on IoT Networks: A Survey," *IEEE Communications Surveys & Tutorials,* Vol. 24, No. 1, 2022, pp. 248–279.

[3] M. El-Hajj, "Leveraging digital twins and intrusion detection systems for enhanced security in IoT-based smart city infrastructures," *Electronics*, Vol. 13, No. 19, 2024, Article No. 3941.

[4] A. Thakkar and R. Lohiya, "A survey on intrusion detection system: feature selection, model, performance measures, application perspective, challenges, and future research directions," *Artificial Intelligence Review*, Vol. 55, No. 1, 2022, pp. 453–563.

[5] G. Rohini, C. Gnana Kousalya and J. Bino, "Intrusion Detection System with an Ensemble Learning and Feature Selection Framework for IoT Networks," *IETE Journal of Research*, Vol. 69, No. 12, 2022, pp. 8859–8875.

[6] H. Zouhri, A. Idri and A. Ratnani, "Evaluating the impact of filter-based feature selection in intrusion detection systems," *International Journal of Information Security*, Vol. 23, No. 2, 2024, pp. 759–785.

[7] M. Fatima, O. Rehman, S. Ali and M. F. Niazi, "ELIDS: Ensemble Feature Selection for Lightweight IDS against DDoS Attacks in Resource-Constrained IoT Environment," *Future Generation Computer Systems*, Vol. 15, 2024, pp. 172–187.

[8] S. Niktabe, "IoT Network Malicious Behaviour Profiling Based on Explainable AI Using LSTM and SHAP," *Master of Applied Science Thesis, Graduate Program in Computer Science*, New York University, Toronto, Canada, 2024.

[9] M. B. Musthafa, S. Huda, Y. Kodera, M. A. Ali, S. Araki, J. Mwaura and Y. Nogami, "Optimizing IoT intrusion detection using balanced class distribution, feature selection, and ensemble machine learning techniques," *Sensors*, Vol. 24, No. 13, 2024, Article No. 4293.

[10] M. A. Soni, M. A. Remli, K. M. Daud and J. A. Amien, "Ensemble learning approach to enhancing binary classification in intrusion detection system for Internet of Things," *International Journal of Electronics and Telecommunications*, Vol. 70, No. 20, 2024, pp. 465–472.

[11] A. Odeh and A. Abu Taleb, "Ensemble-Based Deep Learning Models for Enhancing IoT Intrusion Detection," *Applied Sciences*, Vol. 13, No. 21, 2023, Article No. 11985.

[12] N. V. Chawla, K. W. Bowyer, L. O. Hall and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, Vol. 16, 2002.

[13] Y. Qi, "Random Forest for Bioinformatics Ensemble Machine Learning," *Springer*, US, 2012, pp. 307–323.

[14] M. Sarhan, S. Layeghy and M. Portmann, "Evaluating Standard Feature Sets Towards Increased Generalisability and Explainability of ML-Based Network Intrusion Detection," *Big Data Research*, Vol. 30, 2022.

[15] B. S. Sharmila and R. Nagapadma, "Quantized Autoencoder (QAE) Intrusion Detection System for Anomaly Detection in Resource-Constrained IoT Devices Using RT-IoT2022 Dataset," *Cybersecurity*, 2023.

[16] K. Albulayhi, A. Abu Al-Haija, Q. Alsuhibany, S. A. Jillepalli, M. A. Ashrafuzzaman and F. T. Sheldon, "IoT Intrusion Detection Using Machine Learning with a Novel High Performing Feature Selection Method," *Applied Sciences*, Vol. 12, No. 10, 2022, Article No. 5015.

[17] K. S. Kiran, R. K. Devisetty, N. P. Kalyan, K. Mukundini and R. Karthi, "Building an Intrusion Detection System for IoT Environment Using Machine Learning Techniques," *Procedia Computer Science*, Vol. 171, 2020, pp. 2372–2379.

[18] Q. Abbas, S. Hina, H. Sajjad, K. S. Zaidi and R. Akbar, "Optimization of Predictive Performance of Intrusion Detection System Using Hybrid Ensemble Model for Secure Systems," *PeerJ Computer Science*, Vol. 9, 2023, Article No. e1552.

[19] A. K. Mananayaka and S. S. Chung, "Network Intrusion Detection with Two-Phased Hybrid Ensemble Learning and Automatic Feature Selection," *IEEE Access*, Vol. 11, 2023, pp. 45154–45167.

[20] A. H. Farooqi, S. Akhtar, H. Rahman, T. Sadiq and W. Abbass, "Enhancing Network Intrusion Detection Using an Ensemble Voting Classifier for Internet of Things," *Sensors*, Vol. 24, No. 1, 2023, Article No. 127.

[21] M. Amru, R. J. Kannan, E. N. Ganesh, S. Muthumarilakshmi, K. Padmanaban, J. Jeyapriya and S. Murugan, "Network Intrusion Detection System by Applying Ensemble Model for Smart Home," *International Journal of Electrical & Computer Engineering*, Vol. 14, No. 3, 2024.

[22] V. Shinde, K. Singhal, A. Almogren, V. Dhanawat, V. Karande and A. Rehman, "Ensemble Voting for Enhanced Robustness in DarkNet Traffic Detection," *IEEE Access*, 2024.

[23] W. Yao, L. Hu, Y. Hou and X. Li, "A Lightweight Intelligent Network Intrusion Detection System Using One-Class Autoencoder and Ensemble Learning for IoT," *Sensors*, Vol. 23, 2023, Article No. 4141.

[24] H. A. Al Essa and W. S. Bhaya, "Ensemble Learning Classifiers Hybrid Feature Selection for Enhancing Performance of Intrusion Detection System," *Bulletin of Electrical Engineering and Informatics*, Vol. 13, No. 1, 2024, pp. 665–676.

[25] S. Yaras and M. Dener, "IoT-Based Intrusion Detection System Using New Hybrid Deep Learning Algorithm," *Electronics,* Vol. 13, No. 6, 2024, Article No. 1053.