Vol. 07, No. 3 (2025) 2009-2018, doi: 10.24874/PES07.03B.002

Proceedings on Engineering Sciences

www.pesjournal.net

ENHANCING IOT SECURITY WITH MACHINE LEARNING: A RANDOM FOREST-BASED APPROACH FOR CYBER-ATTACK DETECTION

Ilyas Haman[1]
Zahra Oughannou
Nour El Houda Chaoui
Habiba Chaoui

Received 07.08.2024.
Revised 20.01.2025.
Accepted 10.02.2025.

Keywords:

Internet of Things, IoT security, cyber-attacks, Machine Learning, Deep Learning.

A B S T R A C T

A few industries that have undergone significant changes as a result of the Internet of Things' (IoT) rapid growth are smart cities, manufacturing, and healthcare. But the quick spread of IoT devices has led to serious security flaws, making them vulnerable to several kinds of cyberattacks. This paper presents a robust machine learning methodology employing the Random Forest (RF) model to effectively predict and classify cyber-attacks in IoT networks. Comprehensive data preprocessing techniques were utilized, including data cleaning, feature transformation, balancing using the Synthetic Minority Oversampling Technique (SMOTE), and standard scaling to maintain the integrity and uniformity of data. The suggested method's high accuracy (99%), precision, recall, and F1 score in detecting and classifying cyberattacks were assessed using the CIC-IoT2023 dataset. The results demonstrate how cutting-edge machine learning approaches can improve Internet of Things security and help with the critical challenge of defending IoT networks from sophisticated cyberattacks. This work contributes to the development of more reliable and efficient security solutions for IoT components and establishes the foundation for proactive security actions.

1. INTRODUCTION

Internet of Things (IoT) has grown exponentially in the last few years, transforming a number of industries, including smart cities, manufacturing, transportation, and healthcare (Humayun et al., 2024). Research has concentrated on the Internet of Things (IoT) revolution, which transforms everyday objects into intelligent entities capable of communication with each other and the environment. IoT's main objective is to improve people's lives by cutting expenses, making the best use of public resources, and raising the standard of services offered to individuals (Abdullah et al., 2019). These interconnected devices often lack robust security measures, which renders them vulnerable to various attacks. These flaws can be used by malicious actors to steal confidential information, interfere with operations, or even cause physical harm (Ennafiri & Mazri, 2021; Bel & Sabeen, 2022).

[1] Corresponding author: Ilyas Haman
Email: ilyas.haman@uit.ac.ma

With billions of IoT devices in use today, the possibility for assaults has increased significantly due to the inherent security weaknesses in these networks. With an estimated 125 billion IoT devices by the end of 2030, these settings will serve as a refuge for cybercriminals.

Malicious actors find Internet of Things (IoT) devices appealing because they often have minimal processing and storage capacities, making them more susceptible to attacks like Distributed Denial of Service (DDoS) and data theft.

Firewalls, user authentication (Ayoub, 2020) (Ayoub, 2019), and intrusion detection systems (IDS) are examples of traditional cybersecurity solutions that have not been able to keep up with these new threats (Jullian et al., 2023).

For instance, a recent report by Palo Alto Networks highlights the ongoing threat posed by botnets leveraging compromised IoT devices. In 2023, they observed a significant increase in attacks targeting vulnerabilities in popular smart home devices, leading to large-scale data breaches (Oughannou et al., 2021). This demonstrates how crucial it is to have robust security mechanisms in place in order to protect the increasing number of Internet of Things devices. Traditional security methods may prove inadequate as cyberattacks grow in complexity and sophistication. This paper explores robust methodologies for anomaly detection in IoT environments.

- We introduced a robust machine learning methodology using the Random Forest (RF) model to effectively predict and classify cyber-attacks.
- We improved model performance through detailed data preprocessing techniques, which included data cleaning, feature transformation, Synthetic Minority Oversampling Technique (SMOTE) for balancing, and standard scaling to ensure data integrity and consistency.
- We showcased the efficiency of advanced machine learning methods in enhancing IoT security and tackling the significant challenge of safeguarding IoT networks from sophisticated cyberattacks.

This research makes a significant contribution to the development of more reliable and effective security solutions for Internet of Things (IoT) equipment. By utilizing advanced machine learning techniques to accurately predict and classify cyber-attacks, our approach lays the groundwork for proactive security measures.

The following is how the paper is organized: The Internet of Things is covered in Section 2, relevant literature is reviewed in Section 3, procedures and materials are described in Section 4, results and findings are presented in Section 5, and the study is concluded in Section 6.

2. INTERNET OF THINGS (IOT)

2.1 Definition

Internet of things is a technology enables objects and devices to interact and exchange information autonomously, without human intervention (Abdullah et al., 2019), as illustrated in Figure 1.

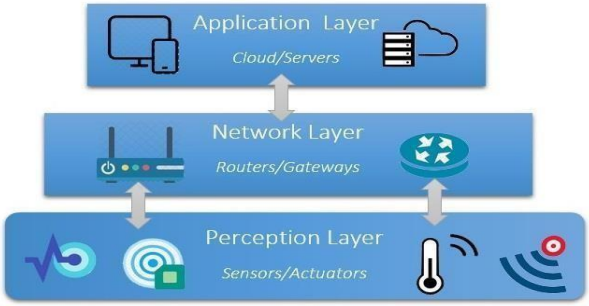

Figure 1. IoT three-layer architecture (Oughannou et al., 2021)

2.2 Architecture

While a universally accepted IoT architecture has yet to be established, a three-layer model with the application layer, network layer, and perceptual or physical layer is proposed by numerous academics (Ennafiri & Mazri, 2021).

Perception layer: The main layer of this function is to spot objects in the IOT ecosystem. This layer is composed of several technologies like sensors or RFID tags (Oughannou et al., 2021).

Network layer: Using gateways, this layer makes it easier for Internet of Things devices to connect directly to the cloud. It processes data obtained from the perception layer for evaluation and decision-making (Ennafiri & Mazri, 2021).

Application layer: This layer functions as the software component within the IoT architecture, responsible for presenting processed data customized for particular application domains like smart cities and healthcare, after the data has been collected and analyzed (Ennafiri & Mazri, 2021).

2.3 Security concern in IOT architecture

While the IoT architecture offers numerous benefits, it also presents security threats at various layers. Table 1 summarizes some of these key threats across each layer.

At the perception layer:

The perception layer encounters numerous security threats and attacks, presenting significant challenges in terms of security.

Eavesdropping: Attackers can intercept IoT dataflow to gather sensitive user information by deploying malicious IoT devices that sniff network traffic (Tewari & Gupta, 2020).

Malicious Data Injection: False sensor data injection refers to the malicious act of forging or altering sensor data utilized in IoT applications, constituting a significant type of attack (Tewari & Gupta, 2020).

Sybil attack: involve malicious nodes assuming multiple identities of legitimate nodes, either by impersonating them or using fabricated identities by replicating. One malicious node can possess multiple identities either simultaneously or one after another (Tewari & Gupta, 2020).

At the network layer:

Different threats have the ability to compromise data availability, confidentiality, and integrity at the network layer of the Internet of Things architecture.

Hello flood: it involves inundating a network or node with numerous route establishment requests. Nodes withing the network may mistake these excessive requests as legitimate communication attempts, thus establishing unnecessary routes and potentially causing congestion or network instability (Tewari & Gupta, 2020).

Traffic analysis: The attacker scrutinizes and stores intercepted traffic for future exploitation. They can then manipulate the interface by replaying previously intercepted traffic intended for the gateway. This scrutinized traffic or data is subsequently repurposed in an altered context (Tewari & Gupta, 2020).

Wormhole: Traffic is intercepted at one place and rerouted to another as part of this network attack. It consequently results in performance problems and network congestion. (Tewari & Gupta, 2020).

In the layer of applications:

The layer of applications, the central hub for managing and utilizing data withing the IoT architecture, is susceptible to various threats and attacks.

Malicious code: Through the internet, malicious code or tailored malware can take advantage of vulnerabilities in Internet of Things devices, giving attackers access to gain access to these devices. Through the compromised devices, they can then initiate additional attacks on other endpoints or networking applications (Tewari & Gupta, 2020).

Cross-site script (XSS): is a tactic used by hackers to infiltrate reputable websites with harmful spyware. If successful, an XSS attack can give the attacker full control over the IoT system (Tewari & Gupta, 2020).

Spyware attack: Installation software installed on IoT devices without permission are designed to collect information, enabling attackers to gather sensitive user data by monitoring their behavior (Tewari & Gupta, 2020).

Table 1. Threats in IOT model

Layer	Threats
Application Layer	Malicious Code
	Cross-Site Script
	Spyware Attack
Network Layer	Hello Flood
	Traffic Analysis
	Wormhole
Perception Layer	Eavesdropping
	Malicious Data Injection
	Sybil Attack

3. RELATED WORK

Machine learning's potential (ML) and deep learning (DL) techniques to enhance IoT security has been thoroughly investigated. Various models have been proposed to detect anomalies, classify threats, and enhance overall system resilience. This section examines a number of studies that have used various ML and DL algorithms to address security concerns in Internet of Things settings.

In (Krishna et al., 2021), the authors propose an SVM-based anomaly detection system for identifying routing layer attacks in IoT networks. The c-SVM model constructs a hyperplane to distinguish between benign and malicious behavior by learning from sensor data gathered during tests under both benign and malignant scenarios. The evaluation, using a "best-case" network topology where the sink node is central, achieved 100% accuracy in detecting malicious nodes. However, the generalizability of this approach to different network configurations is limited, as the accuracy drops to 81% in other scenarios.

The authors in (Ahmad & Alsmadi, 2021) explore the potential of SVMs for detecting HTTP botnets in the IoT environment. The results are promising, with SVMs achieving a high accuracy of 99.86%. However, while achieving perfect recall (100%), the precision of 91.98% indicates a potential for false positives (non-botnet devices misclassified as botnets).

In another paper (Ioannou & Vassiliou, 2019), the effectiveness of Support Vector Machines (SVMs) for cyberattack detection in IoT networks is evaluated. The suggested SVM model demonstrates good results that are accurate of 95.5%, precision of 97.72%, recall of 99.12%, and F1-score of 98.42%, indicating substantial capability for identifying hostile activity within the network.

In (Ioannou & Vassiliou, 2020), To identify selectively forward and blackhole network layer assaults in Internet of Things networks, an intrusion detection model based on SVM was developed by the authors. IoT testbed data that captured both benign and malevolent scenarios made up the dataset used for training and assessment. Network topologies with and without sink-in-the-middle were used to evaluate the SVM model. Findings revealed that while the model achieved a high detection rate of 100% true positive rate and 99.8% accuracy for selective forward and blackhole attacks in optimal network configurations, its precision for detecting selective forward attacks fell below 50%. This discrepancy highlights a limitation in generalizing the model's efficacy across varying network topologies.

In (Krishnan et al., 2021), the authors proposed a method for identifying network attacks in Internet of Things devices that uses many classifiers in supervised learning. The IoTID20 dataset, which logs network activity related to botnet assaults against smart home devices, was utilized by them. Backward Sequential Processing, Sequential Forward Processing, and Recursive Feature Elimination (RFE) feature selection techniques were evaluated in conjunction with the Random Forest, Support Vector Classifier (SVC), and eXtreme Gradient Boosting (XGBoost) algorithms. The results showed that the best accuracy of 99.79% was obtained when combining RFE and XGBoost to distinguish between traffic that is harmful and that is not. The key advantage of this approach is its high detection accuracy across various classifiers and feature selection methods. One of the main drawbacks is that the accuracy of the predictions is very reliant on the caliber of the training data and the features chosen.

In (Nõmm & Bahşi, 2018), with an emphasis on feature selection for dimensionality reduction, an unsupervised anomaly-based detection method was created by the authors to identify botnet attacks in Internet of Things networks. They used traffic from both benign and malware-compromised devices (Bashlite and Mirai) from the N-BaIoT dataset. The study used entropy, variance, and Hopkins statistics for feature selection and evaluated the effectiveness of Local Outlier Factor (LOF), one-class SVM, and isolation forest (IF) for anomaly detection. The highest accuracy of over 90% was achieved using isolation forest with entropy-based feature selection. While the approach effectively reduced computational complexity and provided interpretable results, it faced challenges in capturing the normal behavior of certain IoT devices, leading to lower detection rates for those devices. The key advantage of this method is its efficiency and scalability; however, its generalizability and accuracy across diverse IoT devices can vary.

The authors of (Chaudhary & Gupta, 2019) suggest a framework for identifying DDoS attacks initiated from compromised IoT devices within a local network. This framework uses Random Forest (RF) to analyze traffic and overcome the challenge of limited resources in IoT devices. The model captures benign traffic patterns from legitimate devices, extracts key features, and differentiates between normal and malicious activity. The implemented framework achieved a high accuracy of around 99.17% in detecting DDoS traffic on a Raspberry Pi setup.

Researchers in (Dwyer et al., 2019) assess the effectiveness of different machine learning algorithms for identifying Mirai-alike botnets. The Random Forest classifier significantly outperforms other approaches, achieving impressive accuracy (over 99%), precision (close to 99%), recall (around 98%), and a high F1-score (around 98%).

In (Kumar & Lim, 2019), the authors introduce EDIMA, a modular framework for detecting malicious activity from IoT malware in large networks using K-Nearest Neighbors (k-NN). The evaluation using k-NN achieved an accuracy of 94.44%, demonstrating EDIMA's potential for securing resource-limited IoT devices in large-scale networks.

In (Ullah et al., 2019), The authors suggest using deep learning in two ways to protect networks from malware and software piracy. For malware detection, a deep convolutional neural network (CNN) examines color image visualizations of files from the Maling dataset, achieving a high classification accuracy of 97.46%.

In (Hwang et al., 2019), the researchers present an innovative approach that employs Long Short-Term Memory (LSTM) to categorize harmful material at the packet level. The system uses "word embedding" to capture semantic meaning from packet header fields and to discover the temporal correlations between these domains, attaining a 97.22% accuracy rate.

In another study (Jony & Arnob, 2024), The authors describe a unique method for improving IoT security through intrusion detection that makes use of LSTM networks. Tested on the CIC-IoT2023 dataset, the model showed its effectiveness in recognizing both well-known and new attack patterns, achieving a 98.59% F1 score and a 98.75% accuracy rate.

Researchers in (Altunay & Albayrak, 2023) investigate the potential of combining CNNs and LSTM networks for cyberattack detection in IoT environments. With a high precision of 97.16%, precision of 97.41%, recall of 99.10%, and F1-score of 98.23%, this hybrid technique was successful.

In (Fatayer & Azara, 2019), the authors examine three alternative ANN algorithms: To identify different types of attacks on IoT devices, Multilayer Perceptron (MLP), Radial Basis Function (RBF), and Learning Vector Quantization (LVQ) are used. The MLP has the slowest

processing speed but the best accuracy (99.86%) in threat detection.

The study conducted by(Hodo et al., 2016) investigates the potential application of Artificial Neural Networks (ANNs) in Internet of Things systems for anomaly detection. The evaluation demonstrates that ANNs achieve a high-test accuracy of 99.4%, indicating their effectiveness in identifying abnormal behavior within the network.

4. MATERIALS AND METHODS

4.1. Description of Dataset

This study made use of the CIC IoT Dataset 2023 to facilitate the development of security analytics tools for IoT scenarios. It consists of 33 different attack methods that were used against 105 Internet of Things (IoT) devices and were divided into seven groups: Web-based, Brute Force, Spoofing, Recon, DDoS, and Mirai. These attacks originate from compromised IoT devices, target other IoT devices to disrupt their operations. The dataset provides a thorough summary of these attacks, including those specific to Mirai, such as GREIP Flood and UDPPlain, as well as flooding attacks like UDP and ICMP Floods, SQL Injection, and XSS. Each attack type's frequency is represented by the number of rows in the dataset, highlighting the severity of these threats. The dataset features comprehensive network traffic data, including timestamps, flow duration, protocol types (e.g., IP, UDP, TCP), protocols at the application layer (e.g., HTTP, HTTPS, DNS), data throughput, and packet transmission rates. Statistical metrics like covariance, variance ratio, and weight provide insights into traffic patterns and anomalies, offering a comprehensive understanding of network performance and security.

4.2. Data Preprocessing

Data Cleaning:

In this phase, the dataset undergoes thorough cleaning to ensure data integrity and reliability. This process includes handling missing values by filling them with appropriate statistical measures or using imputation techniques, removing duplicate entries to prevent overfitting, and correcting errors to guarantee that data entries are accurate.

These actions are required to improve the overall quality and dependability of the dataset, which improves the performance and accuracy of the machine learning models employed in the ensuing studies.

Feature Transformation:

In this phase, the dataset's labels are transformed to simplify the classification task. Various types of attacks are grouped into broader categories. For instance, multiple DDoS attack types are labeled as 'DDoS', and other attack types are similarly grouped. This transformation helps in creating a more manageable and understandable dataset.

Figure 2 shows the class distribution before feature transformation, illustrating the original distribution of various specific attack types and benign traffic. After the transformation, Figure 3 shows the class distribution where attack types are grouped into broader categories such as 'DDoS', 'DoS', 'Recon', etc.

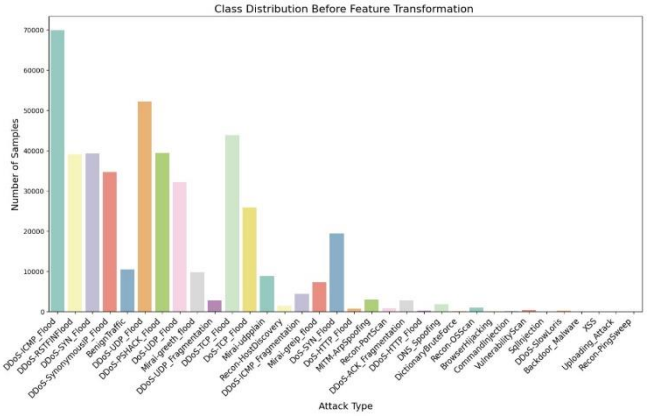

Figure 2. Class Distribution Before Feature Transformation

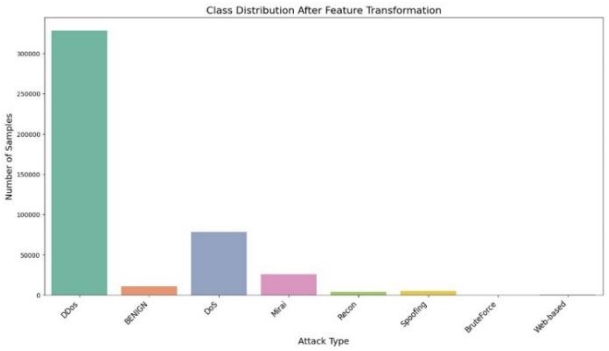

Figure 3. Class Distribution After Feature Transformation

Data balancing using SMOTE

The CIC-IoT2023 dataset exhibits an imbalance between normal and attack traffic, potentially leading to biased model training. To mitigate this issue, we use SMOTE (Synthetic Minority Over-sampling Technique) (Torgo et al., 2013). By interpolating between already-existing data points, SMOTE produces artificial specimens for the minority category, balancing the dataset and enhancing the model's generalization capabilities.

The class distribution prior to applying SMOTE is shown in Figure 4, which highlights the disparity between the typical and different attack classes. After applying SMOTE, Figure 5 displays the distribution of classes, where the classes are more balanced. This balanced

distribution is crucial for effective model training and evaluation.

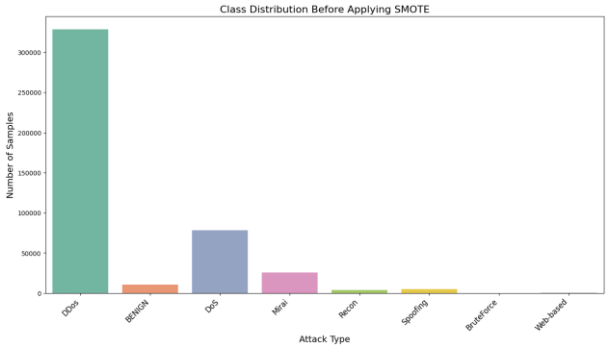

Figure 4. Class Distribution Before Applying SMOTE

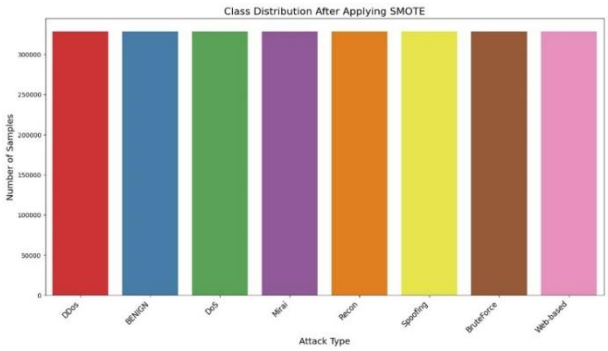

Figure 5. Class Distribution After Applying SMOTE

Scaling using Standard Scaler

In this phase, standard scaling a technique that modifies the characteristics to have a zero mean and a one standard deviation was used. The mean value of each feature is deducted, and the remainder is divided by the standard deviation of the feature. Standard scaling ensures that each feature contributes normalizing the data will contribute equally to the model-training process, as shown in formula (1):

$$X_{Scaled} = \frac{X-\mu}{\sigma} \quad (1)$$

where σ is the standard deviation and μ is the mean of the feature values. Standard scaling is especially important for models that depend on the scale of the input data, such as support vector machines and neural networks. By standardizing the range of the features, we improve the convergence speed of the learning algorithms and enhance the overall accuracy and robustness of the models. This process ensures a more balanced influence of all features during model training, leading to more accurate and reliable predictions.

4.3. Model Architecture

We used a Random Forest (RF) model in our study to identify and classify cyberattacks on Internet of Things networks. During training, the RF model generates several decision trees, indicating the class mode (classification) or mean prediction (regression) for each unique tree. It is a potent machine learning approach. This ensemble approach enhances the model's accuracy and robustness, making it well-suited for the complex task of intrusion detection in IoT environments.

The architecture of our Random Forest model implementation is illustrated in Figure 6. The process begins with the CIC IoT Dataset 2023, which undergoes several preprocessing steps, including data cleaning, feature transformation, SMOTE, and standard scaling. The model is then trained deploying the preprocessed data. The trained model is subsequently tested, with the outputs being the classification of different attack types and benign traffic.

This architecture highlights the comprehensive preprocessing and robust machine learning techniques employed to achieve high-performance cyber-attack detection and classification in IoT networks.

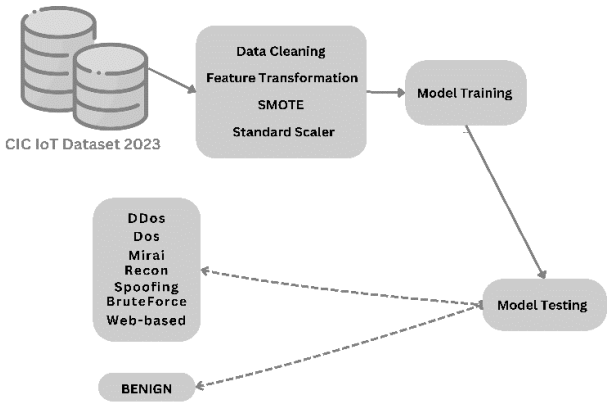

Figure 6. The general architecture of our model

5. RESULTS AND FINDINGS

70% of the dataset was set aside for training and the remaining sets for testing. This distribution was made to maintain consistency and accuracy in experimenting with and evaluating the performance of the Random Forest (RF) model across various data subsets. A large collection of feature columns was established, containing information about several characteristics of network traffic, including protocol types, flow times, counts of flags, includes details regarding protocols such as DNS, HTTP, and HTTPS. To simplify processes related to supervised learning, class labels were to go in the "label" column.

A number of crucial steps were made during data preprocessing to get the dataset ready for instruction and assessment the RF model. NumPy arrays were transformed from the raw data, which included sequences and the labels that associated with them, in order to facilitate further processing. To improve the model's performance, standardization was used to make sure that its attributes had a mean of 0 and a standard deviation of

1. Numerical values were obtained from categorical labels using label encoding, making them suitable for machine learning methods. Due to computational resource limitations, a single CSV file was used instead of the entire dataset. These preprocessing steps ensured that the data was correctly formatted and scaled for optimal model performance, facilitating efficient RF model training and evaluation.

To ensure thorough learning and objective assessment, we divided the dataset into three parts before executing our model: 20% for validation and 70% for training, and 10% for testing. The RF model was trained on this dataset, and its performance was assessed using the F1 score, recall, accuracy, and precision.

With an accuracy of 99.92%, the evaluation results for the RF model were outstanding and showed how precisely and effectively it could classify instances. This high accuracy highlights the model's reliability and practical utility. With an F1 score of 0.9992, recall and precision are traded off in a way that is well-balanced, demonstrating that the model minimizes false positives and false negatives while efficiently detecting significant cases. This strong F1 score is especially important in situations when accuracy and memory are critical.

The recall score of 0.9992 reflects the model's proficiency in capturing actual positive instances, further proving its effectiveness in accurately identifying positive events. Additionally, the precision score of 0.9992 underscores the model's accuracy in correctly classifying objects. Table 2 presents the outcomes of this experiment.

A confusion matrix study was performed to evaluate the efficacy of the RF-based intrusion detection model (refer to Figure 7) was performed on the classification results. This allowed us to assess the accuracy of the model and find any misclassifications. In the confusion matrix, each row represents the actual class labels, and each column represents the expected class labels. The diagonal elements of the matrix indicate true positives and true negatives, or correct predictions, whereas false positives and false negatives, or wrong classifications, are represented by the off-diagonal portions. The number of occurrences in each category is displayed in the matrix elements. The visual representation of the confusion matrix provides valuable information about the accuracy of the model's classification and possible areas for improvement.

The confusion matrix for the RF model demonstrated high accuracy across all classes, with very few misclassifications. This result underscores the robustness of the RF model in detecting and classifying various attack types in IoT network traffic data.

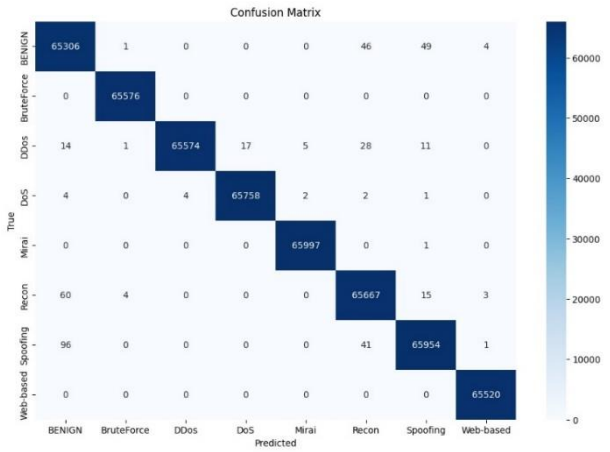

Figure 7. Confusion Matrix

Table 2. Performance metrics of the suggested model

Evaluation matrices	Output
Accuracy	0.9992
F1 Score	0.9992
Recall Score	0.9992
Precision Score	0.9992

6. Conclusion and future work

This study highlights the critical need for strong security measures in the quickly growing IoT device landscape. Because these devices lack reliable security solutions, they are frequently targets of various cyber-attacks. We proposed a comprehensive machine learning methodology using the Random Forest (RF) model to predict and classify cyber-attacks effectively. Through meticulous data preprocessing, including data cleaning, feature transformation, balancing using the Synthetic Minority Oversampling Technique (SMOTE), and standard scaling, we ensured data integrity and uniformity, thereby enhancing the model's performance. The RF model demonstrated exceptional performance, achieving an accuracy of 99.92%, a precision of 0.9992, a recall of 0.9992, and an F1 score of 0.9992. These results underscore the model's reliability and effectiveness in accurately detecting and classifying different forms of cyber-attacks in IoT networks. The confusion matrix further highlighted the model's robustness, with high accuracy across all classes and minimal misclassifications.

While our RF model has shown remarkable results, there are several avenues for future research to further enhance IoT security:

- Examining Alternative Machine Learning Models: To compare and perhaps enhance the outcomes of the RF model, examine the effectiveness of alternative machine learning methods, such as Support Vector Machines (SVM) or Gradient Boosting Machines.
- Deep Learning Approaches: Put into practice and assess deep learning models, like CNNs and Long Short-Term Memory (LSTM) networks,

which may provide better performance in managing the complex and constantly-changing nature of Internet of Things network traffic.

- Real-time Implementation: Develop and test real-time intrusion detection systems that can leverage the trained models for live monitoring and threat IoT network detection.

References:

Abdullah, A., Hamad, R., Abdulrahman, M., Moala, H., & Elkhediri, S. (2019). CyberSecurity: A Review of Internet of Things (IoT) Security Issues, Challenges and Techniques. 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS), 1–6. https://doi.org/10.1109/CAIS.2019.8769560

Ahmad, R., & Alsmadi, I. (2021). Machine learning approaches to IoT security: A systematic literature review. *Internet of Things, 14*, 100365. https://doi.org/10.1016/j.iot.2021.100365

Altunay, H. C., & Albayrak, Z. (2023). A hybrid CNN+LSTM-based intrusion detection system for industrial IoT networks. *Engineering Science and Technology, an International Journal, 38*, 101322. https://doi.org/10.1016/j.jestch.2022.101322

Ayoub, A. N. (2019). Lightweight secure scheme for IOT-cloud convergence based on elliptic curve. *Journal of Theoretical and Applied Information Technology*, 144–155.

Ayoub, A. N. (2020). A lightweight secure CoAP for IoT-cloud paradigm using elliptic-curve cryptography. *Indonesian Journal of Electrical Engineering and Computer Science*, 1460–1470.

Chaudhary, P., & Gupta, B. B. (2019). DDoS Detection Framework in Resource Constrained Internet of Things Domain. 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE), 675–678. https://doi.org/10.1109/GCCE46687.2019.9015465

Dwyer, O. P., Marnerides, A. K., Giotsas, V., & Mursch, T. (2019). Profiling IoT-Based Botnet Traffic Using DNS. 2019 IEEE Global Communications Conference (GLOBECOM), 1–6. https://doi.org/10.1109/GLOBECOM38437.2019.9014300

Ennafiri, M., & Mazri, T. (2021). Internet of Things for Smart Healthcare: A Review on a Potential IOT Based System and Technologies to Control COVID-19 Pandemic. In M. Ben Ahmed, İ. Rakıp Karaș, D. Santos, O. Sergeyeva, & A. A. Boudhir (Eds.), Innovations in Smart Cities Applications Volume 4 (pp. 1256–1269). Springer International Publishing. https://doi.org/10.1007/978-3-030-66840-2_96

Fatayer, T. S., & Azara, M. N. (2019). IoT Secure Communication using ANN Classification Algorithms. 2019 International Conference on Promising Electronic Technologies (ICPET), 142–146. https://doi.org/10.1109/ICPET.2019.00033

Bel, H. F., & Sabeen, S. (2022). A survey on IoT security: attacks, challenges and countermeasures. *Webology*, *19*(1), 3741-3763. https://doi.org/10.14704/WEB/V19I1/WEB19246

Hodo, E., Bellekens, X., Hamilton, A., Dubouilh, P.-L., Iorkyase, E., Tachtatzis, C., & Atkinson, R. (2016). Threat analysis of IoT networks using artificial neural network intrusion detection system. 2016 International Symposium on Networks, Computers and Communications (ISNCC), 1–6. https://doi.org/10.1109/ISNCC.2016.7746067

Humayun, M., Tariq, N., Alfayad, M., Zakwan, M., Alwakid, G., & Assiri, M. (2024). Securing the Internet of Things in Artificial Intelligence Era: A Comprehensive Survey. *IEEE Access, 12*, 25469–25490. https://doi.org/10.1109/ACCESS.2024.3365634

Hwang, R.-H., Peng, M.-C., Nguyen, V.-L., & Chang, Y.-L. (2019). An LSTM-Based Deep Learning Approach for Classifying Malicious Traffic at the Packet Level. *Applied Sciences, 9*(16), Article 16. https://doi.org/10.3390/app9163414

Ioannou, C., & Vassiliou, V. (2019). Classifying Security Attacks in IoT Networks Using Supervised Learning. 2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS), 652–658. https://doi.org/10.1109/DCOSS.2019.00118

Ioannou, C., & Vassiliou, V. (2020). Experimentation with Local Intrusion Detection in IoT Networks Using Supervised Learning. 2020 16th International Conference on Distributed Computing in Sensor Systems (DCOSS), 423–428. https://doi.org/10.1109/DCOSS49796.2020.00073

Jony, A. I., & Arnob, A. K. B. (2024). A long short-term memory based approach for detecting cyber attacks in IoT using CIC-IoT2023 dataset. *Journal of Edge Computing, 3*(1), 28–42. https://doi.org/10.55056/jec.648

Jullian, O., Otero, B., Rodriguez, E., Gutierrez, N., Antona, H., & Canal, R. (2023). Deep-Learning Based Detection for Cyber-Attacks in IoT Networks: A Distributed Attack Detection Framework. *Journal of Network and Systems Management, 31*(2), 33. https://doi.org/10.1007/s10922-023-09722-7

Krishna, R. R., Priyadarshini, A., Jha, A. V., Appasani, B., Srinivasulu, A., & Bizon, N. (2021). State-of-the-Art Review on IoT Threats and Attacks: Taxonomy, Challenges and Solutions. *Sustainability*, *13*(16), Article 16. https://doi.org/10.3390/su13169463

Krishnan, S., Neyaz, A., & Liu, Q. (2021). IoT Network Attack Detection using Supervised Machine Learning. https://hdl.handle.net/20.500.11875/3245

Kumar, A., & Lim, T. J. (2019). EDIMA: Early Detection of IoT Malware Network Activity Using Machine Learning Techniques (arXiv:1906.09715). arXiv. http://arxiv.org/abs/1906.09715

Nõmm, S., & Bahşi, H. (2018). Unsupervised Anomaly Based Botnet Detection in IoT Networks. 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), 1048–1053. https://doi.org/10.1109/ICMLA.2018.00171

Oughannou, Z., Kandrouch, I., Chaoui, N. E. H., & Chaoui, H. (2021). Proposal architecture based on Fog Computing allowing real-time analysis in Smart Cities. 2021 7th International Conference on Optimization and Applications (ICOA), 1–7. https://doi.org/10.1109/ICOA51614.2021.9442671

Tewari, A., & Gupta, B. B. (2020). Security, privacy and trust of different layers in Internet-of-Things (IoTs) framework. *Future Generation Computer Systems*, *108*, 909–920. https://doi.org/10.1016/j.future.2018.04.027

Torgo, L., Ribeiro, R. P., Pfahringer, B., & Branco, P. (2013). SMOTE for Regression. In L. Correia, L. P. Reis, & J. Cascalho (Eds.), *Progress in Artificial Intelligence* (Vol. 8154, pp. 378–389). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-40669-0_33

Ullah, F., Naeem, H., Jabbar, S., Khalid, S., Latif, M. A., Al-turjman, F., & Mostarda, L. (2019). Cyber Security Threats Detection in Internet of Things Using Deep Learning Approach. *IEEE Access*, *7*, 124379–124389. https://doi.org/10.1109/ACCESS.2019.2937347

Ilyas Haman
National School of Applied Sciences
Ibn Tofail University,
Kenitra,
Morocco
ilyas.haman@uit.ac.ma

Zahra Oughannou
National School of Applied Sciences
Ibn Tofail University,
Kenitra,
Morocco
zahra.oughannou@uit.ac.ma
ORCID 0000-0002-6305-2794

Nour El Houda Chaoui
National School of Applied Sciences
Ibn Tofail University,
Kenitra,
Morocco
nourelhouda.chaoui@uit.ac.ma
ORCID 0000-0002-4228-035X

Habiba Chaoui
National School of Applied Sciences
Ibn Tofail University,
Kenitra,
Morocco
habiba.chaoui@uit.ac.ma
ORCID 0000-0001-8892-9612