I.J. Information Technology and Computer Science, **2024, 4, 56-65**
Published Online on August 8, 2024 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijitcs.2024.04.04

Securing the Internet of Things: Evaluating Machine Learning Algorithms for Detecting IoT Cyberattacks Using CIC-IoT2023 Dataset

Akinul Islam Jony*
American International University-Bangladesh (AIUB), Dhaka, 1229, Bangladesh
E-mail: akinul@aiub.edu
ORCID iD: https://orcid.org/0000-0002-2942-6780
*Corresponding author

Arjun Kumar Bose Arnob
American International University-Bangladesh (AIUB), Dhaka, 1229, Bangladesh
E-mail: arjunkumarbosu@gmail.com
ORCID iD: https://orcid.org/0009-0003-2244-2328

Received: 06 November 2023; Revised: 02 January 2024; Accepted: 26 March 2024; Published: 08 August 2024

Abstract: An increase in cyber threats directed at interconnected devices has resulted from the proliferation of the Internet of Things (IoT), which necessitates the implementation of comprehensive defenses against evolving attack vectors. This research investigates the utilization of machine learning (ML) prediction models to identify and defend against cyber-attacks targeting IoT networks. Central emphasis is placed on the thorough examination of the CIC-IoT2023 dataset, an extensive collection comprising a wide range of Distributed Denial of Service (DDoS) assaults on diverse IoT devices. This ensures the utilization of a practical and comprehensive benchmark for assessment. This study develops and compares four distinct machine learning models Logistic Regression (LR), K-Nearest Neighbors (KNN), Decision Tree (DT), and Random Forest (RF) to determine their effectiveness in detecting and preventing cyber threats to the Internet of Things (IoT). The comprehensive assessment incorporates a wide range of performance indicators, such as F1-score, accuracy, precision, and recall. Significantly, the results emphasize the superior performance of DT and RF, demonstrating exceptional accuracy rates of 0.9919 and 0.9916, correspondingly. The models demonstrate an outstanding capability to differentiate between benign and malicious packets, as supported by their high precision, recall, and F1 scores. The precision-recall curves and confusion matrices provide additional evidence that DT and RF are strong contenders in the field of IoT intrusion detection. Additionally, KNN demonstrates a noteworthy accuracy of 0.9380. On the other hand, LR demonstrates the least accuracy with a value of 0.8275, underscoring its inherent incapability to classify threats. In conjunction with the realistic and diverse characteristics of the CIC-IoT2023 dataset, the study's empirical assessments provide invaluable knowledge for determining the most effective machine learning algorithms and fortification strategies to protect IoT infrastructures. Furthermore, this study establishes ground-breaking suggestions for subsequent inquiries, urging the examination of unsupervised learning approaches and the incorporation of deep learning models to decipher complex patterns within IoT networks. These developments have the potential to strengthen cybersecurity protocols for Internet of Things (IoT) ecosystems, reduce the impact of emergent risks, and promote robust defense systems against ever-changing cyber challenges.

Index Terms: Internet of Things, Cybersecurity, Machine Learning, DDoS Attacks, CIC-IoT2023 Dataset.

1. Introduction

The IoT has become a crucial aspect of our daily lives, and because of its expanding use, there has been a rising number of cyberattacks on IoT devices. Security professionals and academics are extremely concerned about the current situation of IoT cyberattacks. IoT device threats fall under several areas, including network assaults, software attacks, and physical attacks. Node cloning attacks are one type of physical assault that allows for node replication and network access [1]. Advanced Persistent Threat (APT) assaults on software are one type of attack that allows attackers to access a system while going lengthy periods [2]. Attackers can overwhelm a network with traffic and bring it to a halt using DDoS assaults

[3]. IoT device security and privacy are significant problems, and poor authorization and authentication can result in privacy issues at the device level [4]. IoT device vulnerabilities and threats are growing daily, therefore it's critical to create strong defenses to keep them safe. Encryption, authentication, and access control are a few of the countermeasures [5]. It's critical to keep up with the most recent security techniques and technological advancements to prevent assaults on IoT devices.

Cyberattacks on the IoT are becoming more common, and their frequency is rising. According to [6], the overall average number of weekly attacks on IoT devices per business increased by 41% in the first two months of 2023 compared to 2022. The most often targeted IoT devices are those found in European businesses, with APAC and Latin American-based corporations following behind. On average, 54% of organizations experience attempted cyber-attacks every week. IoT device threats can be divided into many types, including network assaults, software attacks, and physical attacks [7]. Node cloning attacks are one type of physical assault that allows for node replication and network access. APT assaults are one type of software attack where an attacker can enter a system and be undiscovered for a lengthy period [8]. DDoS assaults on networks can overwhelm the system with traffic and bring it to a halt. The creation of efficient defenses against cyberattacks, such as access control, authentication, and encryption, is crucial [9].

As a result of these increasing concerns, machine learning (ML) algorithms have surfaced as crucial instruments in proactively identifying and mitigating cyber threats in Internet of Things (IoT) ecosystems. By capitalizing on pre-existing datasets and conducting statistical analysis, machine learning techniques have demonstrated their capacity to detect threats early, identify network vulnerabilities, and decrease operational expenses [10, 11]. Despite these developments, a definitive benchmark for the most effective machine learning algorithms to detect IoT cyber threats has yet to be established, creating a significant void in the field of IoT cybersecurity research [6, 7]. A report on ML-based identification of malware in executable files claims that ML techniques have been used to solve a variety of worldwide computer security issues, including intrusion detection, fraud detection, ransomware recognition, and malware detection [12]. To prevent cyberattacks on IoT devices, it is vital to stay up to speed with the most recent technologies and approaches. ML algorithms are used in cybersecurity to detect and mitigate cyberattacks [13].

This research endeavors to fill this critical void by examining the construction and comparison of machine learning prediction models that employ the CIC-IoT2023 dataset to identify intrusions targeting IoT devices. The dataset presents a practical standard that includes a wide range of DDoS attacks on different IoT devices, thereby offering a broad spectrum for assessing the effectiveness of ML algorithms in the context of IoT cybersecurity [14]. Logistic Regression (LR), K-Nearest Neighbors (KNN), Decision Tree (DT), and Random Forest (RF) are the ML models under consideration. The principal aim is to determine the most efficient machine learning methodologies customized for Internet of Things (IoT) security. This will furnish researchers and practitioners with indispensable knowledge to strengthen their defenses against emergent cyber threats. As far in terms of identifying and reducing evolving attack vectors is concerned, current methodologies and techniques have demonstrated their limitations in the face of growing apprehensions regarding IoT cyber threats. Conventional methodologies frequently confront the complex and varied characteristics of IoT networks, resulting in intrinsic deficiencies when it comes to precisely detecting and averting advanced cyber threats.

The other sections of the paper are organized as follows: section 2 describes the relevant literature, section 3 outlines the methodologies and materials employed in this study, section 4 analyzes the findings, and section 5 offers a concluding summary of the study.

2. Related Works

Extensive research has been conducted in the past to investigate the application of Machine Learning (ML) and Deep Learning (DL) methods to IoT cybersecurity. Nevertheless, these methodologies frequently encounter obstacles when attempting to manage the intricate and ever-changing characteristics of cyber threats that are specifically targeted at interconnected IoT devices. Prevalently flawed are established methodologies, particularly concerning their capacity to thoroughly detect and thwart innovative attack vectors that exploit susceptibilities across heterogeneous IoT ecosystems. The IoT is a rapidly growing industry that permeates everyday existence. Because IoT devices are networked, they are susceptible to cyberattacks. The number of cyberattacks on IoT systems has increased recently, thus it's critical to recognize and prepare for these attacks. Nowadays, it is very common to apply DL and ML-based algorithms as possible solutions to this problem. Consequently, this study will investigate the findings of current research on the use of ML and DL methods for identifying and predicting cyberattacks on IoT devices. The research in [15] conducted a survey and a literature analysis on ML and DL methods for IoT security. To assess how well different ML-based algorithms performed, they used the KDD-99 dataset. The study discovered that cyberattack detection in IoT systems may be accomplished using both ML and DL techniques. The authors also emphasized the need for additional studies to enhance the precision and effectiveness of these approaches. In [16], the use of ML and data analytics for IoT security is covered using random forests, decision trees, and neural networks. They used the NSL-KDD dataset, and the accuracy rate of the RF technique was 99.6%.

ML algorithms are suggested for automating the detection of cyberattacks as well as for quick prediction and analysis of attack types [17]. A deep learning methodology is suggested in another study [18] for anticipating cybersecurity assaults on the IoT. The study uses ML and DL methods to carefully extract important information from a BoT dataset. They showed the improved accuracy performance and dependability of cyber threat prediction in IoT scenarios. The study

produces more precise and reliable forecasts and enhanced IoT security. In a survey [15] of ML and DL techniques for assessing cybersecurity in IoT, various ML techniques are explored for anomalous activities and cyber threats detection using the KDD-99 dataset.

The Bot-IoT dataset [19] is made up of simulated IoT sensor data that includes both normal and attack traffic. Using ML and DL models, an intrusion detection system (IDS) was created to identify the class imbalance issue of the dataset. The DT and multi-layer perceptron models outperformed all other models in the performance evaluation of different models employing three distinct feature sets for identifying DDoS and DoS assaults across IoT networks. More than 99% accuracy on average. The study also showed that, for future Bot-IoT dataset implementations, the Argus flow data generator is not required. ML approaches were used by [20] to create the best security models for spotting IoT intrusions. They used the N-BaIoT dataset, which comprises botnet attacks injected into various IoT devices such as doorbells, baby monitors, security cameras, and webcams, and they primarily focused on botnet attacks targeting different IoT devices. They use a variety of ML models, including deep learning models, in their botnet detection algorithms for each device. With a focus on the models that attained a high detection F1-score, the effectiveness of the models was examined through multiclass and binary classification. The findings demonstrated that ML-based models, in particular deep learning models, were successful in identifying botnet attacks on IoT devices. The findings revealed how ML techniques enhance IoT security and solve issues brought on by the proliferation of IoT devices and threats.

For IoT systems, [21] suggests a paradigm for the next-generation cyber-attack prediction that uses the CHAID decision tree and multi-class SVM to predict cyberattacks with a 99.72% accuracy rate. To detect cyberattacks in IoT networks, [22] presents a DL-based detection method. The study uses LSTM to identify network intrusions and focuses on the detection of DDoS attacks. The study achieves great accuracy rates in complicated assault detection and prediction. The article covers the deep learning models, datasets, and distributed attack detection systems that were created. The research evaluates the distributed attack detection framework and demonstrates the efficacy of distributed DL models to enable IoT networks to detect a wide range of assaults with high detection and accuracy rates.

3. Methods and Materials

When choosing the ML models for this research, we considered the inherent deficiencies of traditional methods when it came to identifying and addressing emergent cyber threats in IoT networks. The inability of conventional approaches to handle the ever-changing, varied, and dynamic characteristics of attack vectors served as the impetus for our investigation into more resilient models that could discern complex patterns in IoT traffic. Similarly, the selection of evaluation metrics was influenced by the deficiencies identified in previous evaluations, intending to rectify the issues and offer a holistic assessment of the model's efficacy that extended beyond traditional metrics. The dataset, ML models, and assessment measures that we employed in this study are also covered in detail. Fig 1 depicts the overall workflow of our technique. Working with the CIC-IoT2023 dataset requires following a prescribed procedure. Loading the dataset is the first step, followed by the essential stage of data preprocessing, which involves handling missing values, cleaning the data, and formatting modifications. Then, to make training and evaluating models easier, the dataset is divided into two subsets- training, and testing. ML techniques are then assessed on the testing set to determine their performance after being selected and trained on the training set. A detailed evaluation of the models' efficacy is conducted using relevant measures, including F1 score, accuracy, precision, and recall. The end goal is to choose the model that best fits the requirements of the ongoing project or to consider further optimization for improved accuracy. This methodical approach for working using the CIC-IoT2023 dataset is ensured by this well-organized methodology, leading to intelligent decisions and reliable ML outcomes.

3.1. Dataset Overview

The CIC-IoT2023 dataset [14], a publicly available dataset that contains actual network traffic from various IoT devices under both normal and attack circumstances, is the one that we employed in this study. The Canadian Institute for Cybersecurity (CIC) and the Information Technology University of Copenhagen (ITU) collaborated to generate the CIC-IoT2023 dataset. A smart home environment with 20 IoT gadgets, including cameras, thermostats, smart TVs, smart watches, etc., was simulated to create the dataset. Wireshark and TCPdump tools were used to record the network traffic, while Snort and Suricata intrusion detection systems were used to categorize it. Ten days' worth of network traffic—five days of regular traffic and five days of attack traffic-make up the dataset. TCP SYN Flood, UDP Flood, HTTP Flood, HTTP Slow Post, Slowloris, MQTT Flood, CoAP Flood, WS-DDoS (WebSocket), Web Service Flood (SOAP), and Web Service Flood (RESTful) are among the ten various DDoS attack types included in the dataset. There are around 80 million packets in the dataset, 64 million of which are classified as malicious and 16 million as normal. For each packet in the dataset, there are 115 features, including the protocol, payload size, timestamp, and source and destination IP addresses.

Fig 2 shows how different cyberattacks are distributed in several instances in the dataset. The graphic deftly classifies fewer common attacks into an "Other" category while highlighting the frequency of various attack kinds. The "Other" category is utilized when the quantity of occurrences for a specific attack is less than a predetermined threshold. This method offers a concise summary of the most common attack routes without overcrowding the chart with labels.

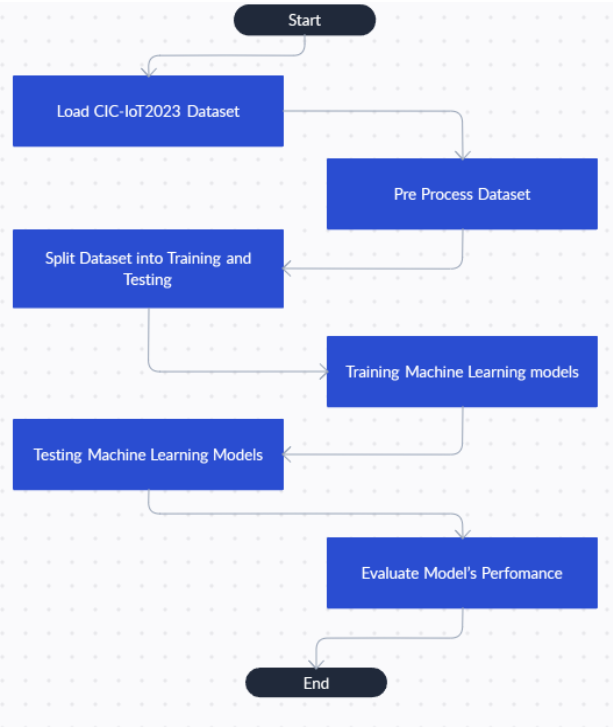

Fig.1. Architecture model of machine learning approach

This dataset differs from other IoT datasets used in network intrusion detection studies in that it possesses the following features:

- Instead of simulating or emulating devices, it uses actual IoT devices as both attackers and victims.
- In contrast to a small number of devices from a single vendor or protocol, it encompasses a broad variety of IoT devices from several manufacturers and protocols.
- Instead of a single type of attack that targets a particular layer or service, it consists of various DDoS attack types that target various layers of the network stack.
- Instead of a small amount of data with low diversity and complexity, it offers a vast amount of data with great diversity.

This dataset can offer a more complex and realistic environment for testing how well ML algorithms work for identifying IoT cyberattacks.

3.2. Machine Learning Models

Using the CIC-IoT2023 dataset, we selected and compared four well-known machine-learning algorithms: RF, DT, KNN, and LR. These algorithms were picked based on how well-liked and effective they were in earlier research on network intrusion detection. With the help of the Python and scikit-learn libraries, we developed these algorithms. Except for KNN, where we changed the number of neighbors to 5, we used the default settings for each algorithm's parameters. Before supplying the dataset to the ML models, we also performed certain preprocessing operations on it. These actions comprise:

- Removing features like packet ID, checksum, and other unused or superfluous components.
- Converting categorical characteristics, such as protocol type and service type, into numerical values.
- Using min-max scaling, numerical features are normalized into a range of [0, 1].
- Using the random under-sampling technique, one can equalize the class distribution by lowering the number of malicious packets to the same level as the number of legitimate packets.
- Dividing the dataset, keeping the class proportion constant, into a training set (70%) and a testing set (30%) using a stratified sampling approach.

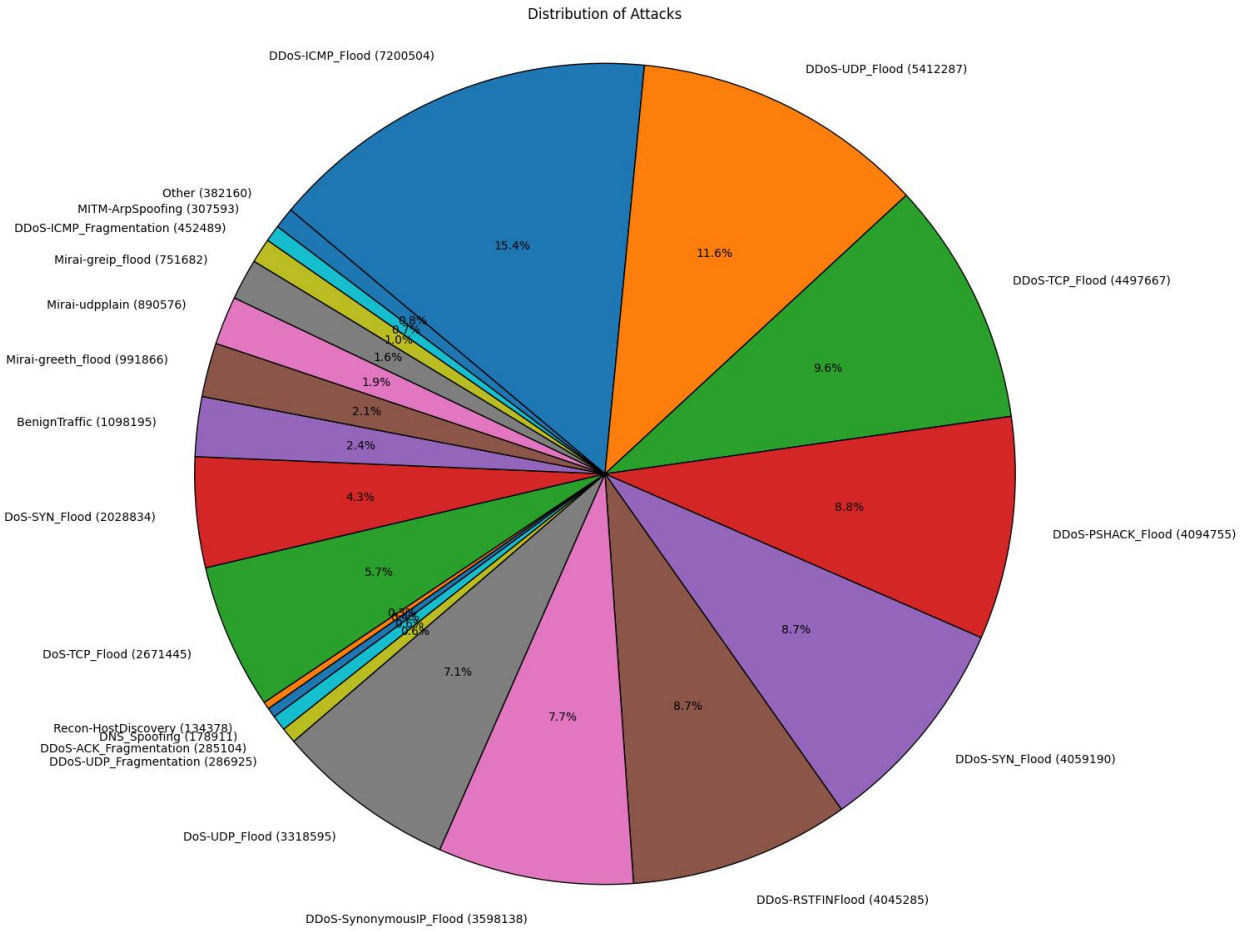

Fig.2. Distribution of attacks

3.3. Evaluation Metrics

On the CIC-IoT2023 dataset, we assessed the evaluation of the ML algorithms using various metrics that are frequently employed in classification tasks. The most common evaluation metrics are accuracy, precision, recall, and F1-Score which are briefly described below along with the equation to calculate.

- Accuracy: The proportion of correctly categorized packets to all packets.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

- Precision: The proportion of harmful packets accurately identified relative to all malicious packets expected.

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

- Recall: The proportion of harmful packets that were accurately identified to all malicious packets.

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

- F1-Score: The harmonic means of the recall and precision.

$$F1 - Score = \frac{Precision+Recall}{2} \quad (4)$$

4. Results and Discussion

To identify IoT cyberattacks, four ML models have been developed using RF, KNN, DT, and LR algorithms. The performance assessment of these models with precision-recall curves is displayed in Fig 3 for each of these algorithms.

A precision-recall curve is a graph that shows the trade-off between precision and recall at different probability thresholds. Precision is the percentage of accurate positive predictions, whereas recall is the proportion of positive incidents that were correctly predicted. The curve of a perfect model would reach the top right corner, signifying 100% recall and 100% precision. The model's performance across all thresholds is gauged by the area under the curve (AUC). DT and RF have the highest AUC, followed by KNN and LR, as can be shown. Due to their ability to distinguish between the most hostile and legitimate packets, DT and RF are therefore the most accurate and trustworthy models for detecting IoT intrusions. While KNN also works well, its precision is lower than that of DT and RF. The algorithm with the lowest AUC, LR, is unsuitable for this task due to its high rate of false positives and false negatives.

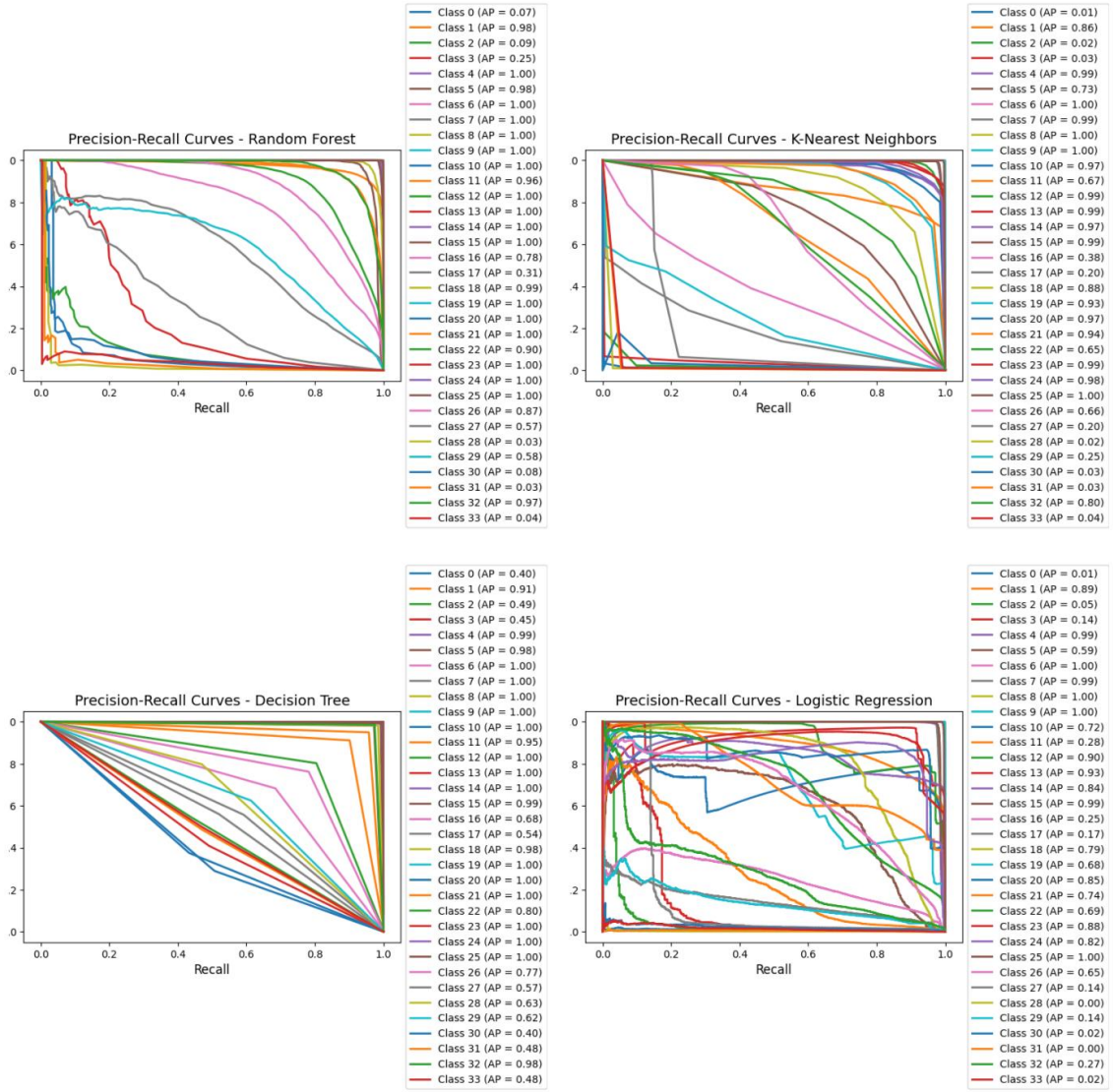

Fig.3. Precision-recall curves

The confusion matrix for each of these ML-based models is displayed in Fig 4, 5, 6, 7. The rows depict the actual classes, while the columns display the predicted classes. The diagonal elements show the correct forecasts. On the other hand, off-diagonal elements show incorrect forecasts. The confusion matrix can be used to calculate a variety of metrics, such as recall, precision, accuracy, and F1-score. In contrast to false positives (FP) and false negatives (FN), which are at their lowest levels, the percentage of true positives (TP) and true negatives (TN) is highest for DT and RF. This indicates that they have a low mistake rate and can correctly identify the majority of packets as malicious or legitimate. KNN has a lot of TP and TN as well, but it also has more FP and FN than DT and RF. This indicates that it has a greater error rate and that some packets may be mistakenly classified as harmful or legitimate. The proportion of TP and TN is lowest while the proportion of FP and FN is largest in LR. This indicates that it has an extremely high error rate and can rarely distinguish between malicious and legitimate messages.

The outcomes highlight how crucial it is to pick the best ML algorithm for IoT threat detection. Some techniques, such as feature selection, dimensionality reduction, parameter tuning, and ensemble methods, can be used to improve the evaluation of ML algorithms. These techniques can maximize the algorithms' potential and raise their effectiveness in spotting IoT assaults.

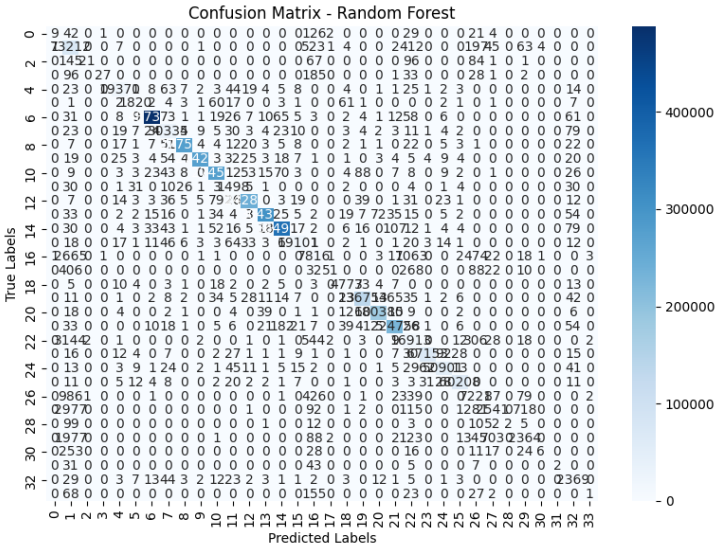

Fig.4. Confusion matrix (random forest)

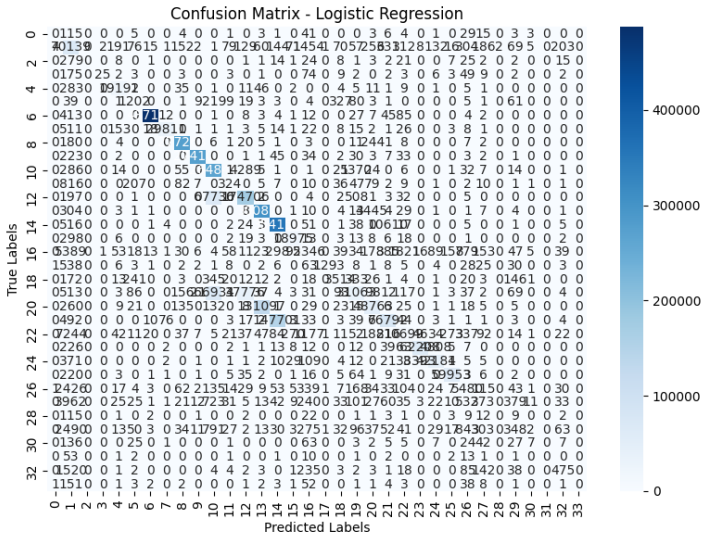

Fig.5. Confusion matrix (logistic regression)

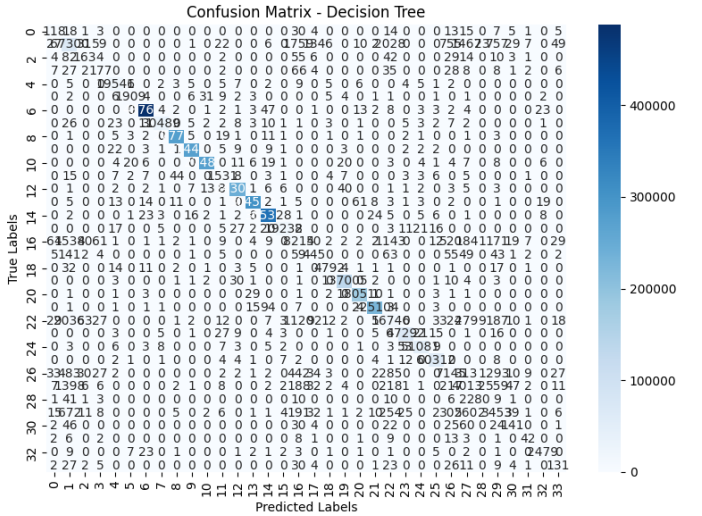

Fig.6. Confusion matrix (decision tree)

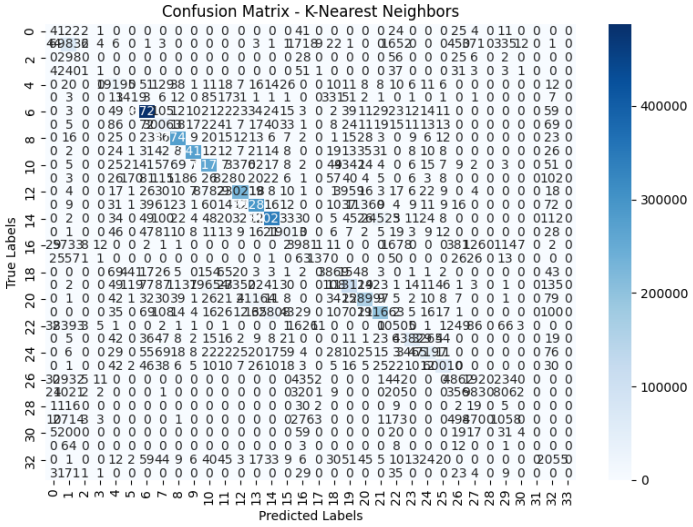

Fig.7. Confusion matrix (k-nearest neighbors)

The performance evaluations of each of these models on the CIC-IoT2023 dataset for detecting cyber-attacks are presented in Table 1. The evaluation metrics include accuracy, precision, recall, and F1 score for comparing these techniques based on RF, KNN, DT, and LR algorithms.

We can make several significant conclusions and observations on the performance of ML algorithms in identifying IoT cyberattacks based on the evaluation metrics that the algorithms DT and RF excelled, attaining the highest accuracy ratings of 0.9919 and 0.9916, respectively. Additionally, they earned the highest precision, recall, and F1-score, showing that they are reliable in correctly categorizing both valid and malicious packets.

With an accuracy of 0.9380, KNN performed reassuringly, effectively and efficiently. Although it may not be as accurate as DT and RF, KNN shows proficiency with non-linear data. KNN can, however, be computationally expensive and is sensitive to noise and outliers. With a score of 0.8275, Logistic Regression (LR) had the lowest accuracy of all the methods. This may be explained by the linear assumption made by LR, which leads to a high percentage of false positives. Additionally, it has low precision and F1-score values due to its sensitivity to noise and outliers. The most successful and efficient algorithms for identifying IoT cyberattacks were Decision Tree and Random Forest. These results offer important information for choosing the best ML algorithms and defense tactics to protect the IoT from online threats.

Table 1. Evaluation metrics of the ML models

Algorithm	Accuracy	Precision	Recall	F1-Score
RF	0.9916	0.9913	0.9916	0.9909
KNN	0.9380	0.9366	0.9380	0.9364
DT	0.9919	0.9920	0.9919	0.9919
LR	0.8275	0.8473	0.8275	0.8034

A detailed evaluation of the effectiveness of machine learning algorithms in mitigating cyber threats to the Internet of Things was conducted through our analysis of the CIC-IoT2023 dataset. Upon examining the precise evaluation metrics associated with each model, it was observed that DT and RF exhibited outstanding performance. RF demonstrated a remarkable F1 Score of 99.08%, recall of 99.15%, and precision of 99.12%, in addition to an accuracy rate of 99.15%. In a similar vein, DT demonstrated exceptional performance with an accuracy of 99.18%, an F1 Score of 99.19%, a recall of 99.18%, and a precision of 99.19%. The robustness of both models in differentiating benign from malevolent packets in IoT networks is highlighted by these metrics. On the other hand, the KNN algorithm demonstrated a noteworthy accuracy of 93.81%. This was supported by F1 Score, recall, and precision values of 93.64%, 93.81%, and 93.66%, respectively. In contrast, LR performed less effectively, achieving an accuracy of 82.75%. This resulted in comparatively lower values for F1 Score, recall, and precision, which were 80.34%, 82.75%, and 84.73% respectively. This detailed examination is consistent with our research goals, as it clarifies the intricate functioning of each model and provides evidence for the superiority of DT and RF in strengthening IoT cybersecurity.

5. Conclusions

In this study, we analyzed four ML algorithms for detecting IoT cyberattacks using the CIC-IoT2023 dataset: RF, KNN, DT, and LR. The dataset used in this study provides a comprehensive and realistic benchmark containing multiple types of DDoS attacks on different IoT devices. We carried out data preparation, model training, and performance

evaluation using relevant metrics like accuracy, precision, recall, and F1-score. The results show that DT and RF are the most successful and efficient algorithms for identifying IoT cyberattacks, with accuracy rates of 0.9919 and 0.9916, respectively. These algorithms are also the best in terms of precision, recall, and F1-score values, indicating that they can reliably distinguish between malicious and normal packets. With an accuracy of 0.9380, KNN does admirably as well, while LR has the lowest accuracy at 0.8275.

This study provides a substantial critique of the inherent constraints that exist in existing approaches to IoT cybersecurity. Through a comprehensive examination of the effectiveness of machine learning models in detecting cyber threats in the context of the Internet of Things (IoT) and utilizing the CIC-IoT2023 dataset, this research sheds light on the limitations of conventional methods in managing the ever-changing and intricate nature of such threats. The potential of DT, and RF algorithms to rectify these shortcomings is highlighted by their superior performance. This could result in a more dependable and efficient method of detecting and thwarting malicious activities in interconnected IoT environments. The results of this research have substantial ramifications for the pragmatic implementation of machine learning in fortifying the security of the Internet of Things. The algorithmic capabilities of DT, and RF demonstrate exceptional levels of accuracy, precision, and recall, rendering them feasible contenders for prompt implementation in IoT defense systems. The ability of these systems to differentiate between benign and malicious traffic provides a strong basis for developing strategies to detect and mitigate threats in real time. This provides concrete advantages for industry stakeholders who are interested in protecting IoT ecosystems.

Moreover, the study establishes a foundation for numerous expansions and forthcoming trajectories in the realm of IoT cybersecurity. Investigating federated learning techniques, incorporating unsupervised learning approaches, and integrating deep learning models are all potentially fruitful avenues for improving the scalability and adaptability of cyber threat detection mechanisms in IoT networks. Moreover, to address the ever-changing cyber threat landscape, enhancing the security of IoT infrastructures could be accomplished through the integration of machine learning algorithms that are continuously improved and diverse datasets are utilized.

Dataset Availability Statement

The dataset used in this study can be found on https://www.unb.ca/cic/datasets/iotdataset-2023.html, [accessed on 05 October 2023].

References

[1] U. Tariq, I. Ahmed, A. K. Bashir, K. Shaukat, "A Critical Cybersecurity Analysis and Future Research Directions for the Internet of Things: A Comprehensive Review". Sensors, Vol. 23, No. 8, 2023. DOI: https://doi.org/10.3390/s23084117

[2] X. Cheng, J. Zhang, B. Chen, "Cyber Situation Comprehension for IoT Systems based on APT Alerts and Logs Correlation", Sensors, Vol.19, No.18, 2019. DOI: https://doi.org/10.3390/s19184045

[3] P. K. Sadhu, V. P. Yanambaka, A. Abdelgawad, "Internet of Things: Security and Solutions Survey", Sensors, Vol. 22, No. 19, 2022. DOI: https://doi.org/10.3390/s22197433

[4] S. Kumar, P. Tiwari, M. Zymbler, "Internet of Things is a revolutionary approach for future technology enhancement: a review", Journal of Big Data, Vol.6, No.1, pp.1-21, 2019. DOI: https://doi.org/10.1186/s40537-019-0268-2

[5] J. P. A. Yaacoub, H. N. Noura, O. Salman, A. Chehab, "Robotics cyber security: Vulnerabilities, attacks, countermeasures, and recommendations", International Journal of Information Security, Vol.21, pp.115-158, 2022. DOI: https://doi.org/10.1007/s10207-021-00545-8

[6] Check Point Research, "The Tipping Point: Exploring the Surge in IoT Cyberattacks Globally", 2023. Retrieved on October 12, 2023, from https://blog.checkpoint.com/security/the-tipping-point-exploring-the-surge-in-iot-cyberattacks-plaguing-the-education-sector/

[7] K. Tsiknas, D. Taketzis, K. Demertzis, C. Skianis, "Cyber threats to industrial IoT: a survey on attacks and countermeasures", IoT, Vol. 2, No. 1, pp. 163-186, 2021. DOI: https://doi.org/10.3390/iot2010009

[8] M. Abdullahi, Y. Baashar, H. Alhussian, A. Alwadain, N. Aziz, L. F. Capretz, S. J. Abdulkadir, "Detecting cybersecurity attacks in the internet of things using artificial intelligence methods: A systematic literature review", Electronics, Vol. 11, No. 2, 2022. DOI: https://doi.org/10.3390/electronics11020198

[9] Ani Petrosyan, "Annual number of IoT attacks global 2022", 2023. Retrieved on October 12, 2023 https://www.statista.com/statistics/1377569/worldwide-annual-internet-of-things-attacks/

[10] M. Ahsan, K. E. Nygard, R. Gomes, M. M. Chowdhury, N. Rifat, J. F. Connolly, "Cybersecurity threats and their mitigation approaches using Machine Learning—A Review", Journal of Cybersecurity and Privacy, Vol. 2, No. 3, pp. 527-555, 2022.

[11] Matthew Urwin, "Machine Learning in Cybersecurity: How It Works and Companies to Know", 2023. Retrieved on October 12, 2023 https://builtin.com/artificial-intelligence/machine-learning-cybersecurity

[12] J. Singh, J. Singh, "A survey on machine learning-based malware detection in executable files", Journal of Systems Architecture, Vol. 112, 2021.

[13] N. Vadivelan, K. Bhargavi, S. Kodati, M. Nalini, "Detection of cyber-attacks using machine learning. In AIP Conference Proceedings." AIP Publishing. Vol. 2405, No. 1, 2022.

[14] E. C. P. Neto, S. Dadkhah, R. Ferreira, A. Zohourian, R. Lu, A. A. Ghorbani, "CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment", Sensors, Vol. 23, No. 13, 2023. DOI: https://doi.org/10.3390/s23135941

[15] U. Inayat, M. F. Zia, S. Mahmood, H. M. Khalid, M. Benbouzid, "Learning-based methods for cyber-attack detection in IoT systems: A survey on methods, analysis, and future prospects", Electronics, Vol. 11, No. 9, 2022.

[16] E. Adi, A. Anwar, Z. Baig, S. Zeadally, "Machine learning and data analytics for the IoT", Neural computing and applications,

Vol. 32, pp. 16205-16233, 2020.
[17] C. Malathi, I. N. Padmaja, "Identification of cyber-attacks using machine learning in smart IoT networks", Materials Today: Proceedings, Vol. 80, pp. 2518-2523, 2023.
[18] O. A. Alkhudaydi, M. Krichen, A. D. Alghamdi, "A Deep Learning Methodology for Predicting Cybersecurity Attacks on the Internet of Things. Information", Vol. 14, No. 10, pp. 550, 2023.
[19] J. G. Almaraz-Rivera, J. A. Perez-Diaz, J. A. Cantoral-Ceballos, "Transport and application layer DDoS attacks detection to IoT devices by using machine learning and deep learning models", Sensors, Vol. 22, No. 9, 2022.
[20] J. Kim, M. Shim, S. Hong, Y. Shin, E. Choi, E. "Intelligent detection of IoT botnets using machine learning and deep learning", Applied Sciences, Vol. 10, No. 19, 2023.
[21] S. Dalal, U. K. Lilhore, N. Foujdar, S. Simaiya, M. Ayadi, N. A. Almujally, A. Ksibi, "Next-generation cyber-attack prediction for IoT systems: leveraging multi-class SVM and optimized CHAID decision tree", Journal of Cloud Computing, Vol. 12, No. 1, pp. 1-20, 2023.
[22] O. Jullian, B. Otero, E. Rodriguez, N. Gutierrez, H. Antona, R. Canal, "Deep-Learning Based Detection for Cyber-Attacks in IoT Networks: A Distributed Attack Detection Framework", Journal of Network and Systems Management, Vol. 31, No. 2, pp. 33, 2023.

Authors' Profiles

Dr. Akinul Islam Jony currently holds the position of Associate Professor and serves as the Head of the Undergraduate Program in Computer Science at American International University-Bangladesh (AIUB). His research interests encompass a wide range of topics, including cybersecurity, artificial intelligence, machine learning, e-learning, educational technology, and issues in software engineering.

Arjun Kumar Bose Arnob is a final semester student of BSc in Computer Science and Engineering and majoring in Software Engineering at the American International University-Bangladesh (AIUB). He is currently working as a Research Assistant at AIUB and is actively involved in research projects. He has a strong passion and proficiency in Machine Learning and Deep Learning which is reflected in his work. He has consistently performed well academically and is dedicated to his studies.

How to cite this paper: Akinul Islam Jony, Arjun Kumar Bose Arnob, "Securing the Internet of Things: Evaluating Machine Learning Algorithms for Detecting IoT Cyberattacks Using CIC-IoT2023 Dataset", International Journal of Information Technology and Computer Science(IJITCS), Vol.16, No.4, pp.56-65, 2024. DOI:10.5815/ijitcs.2024.04.04