Kanak al. International Journal of Recent Research Aspects ISSN: 2349-7688, Special Issue: International Conference on Intelligent Systems for Complex, Chaotic, & Connected Environments Sept 2025, pp. 114-128

Detection And Prediction of IOT Cyber Network Attacks Using Machine Learning

[1]Kanak Ahlawat, [2]Keerti Thakur
Sagar Institute of Research and Technology, Bhopal,India
Maharishi University of Information Technology, Noida, India
Kanakahlawat30@gmail.com
Keertitha20@gmail.com

ABSTRACT— **Progress in machine learning models is fundamentally changing the security protocols for IoT systems, promoting protection measures as the means in which connected smart devices (i.e. sensors, cameras, home automation systems, etc.) will produce large amounts of data for the ML models to study/analyze for unusual or anomalous behavior, therefore enabling the ability to pre detect cyberattacks. This enhances security systems to take a proactive stance, enabling the system to take a proactive role to take action before the attack occurs i.e. identify a hacked device or identify malware that exists in the IoT infrastructure. However, at the same time, these new advanced technologies are being used by malicious actors to bypass security systems and evade defense mechanisms. The hacker will now use these tools as part of their toolkit to devise and circumvent the security strategies, by finding vulnerabilities within devices and exploiting them. Therefore, although machine learning techniques advanced the IoT security environment, it also creates risk. The challenge is that the risk is now balanced against legitimate and ethical practices of actually applied technology implementation.**

Keywords— IoT Cybersecurity, Machine Learning, Cyberattacks, Anomaly Detection, Threat Prediction, Network Defense, Vulnerability Exploitation, Adversarial Machine Learning.

I. INTRODUCTION

Over the past few years, the Internet of Things (IoT) has emerged extremely rapidly as it connects more devices in various industry sectors to create more reliable and efficient human outcomes. However, this rapid expansion has also resulted in an increase in cyberattack risk to the IoT landscape, as many of the IoT devices do not offer very good security. Therefore, it would be beneficial to use machine learning (ML) to improve IoT smartness and security. ML entails using advanced computer algorithms to empower the IoT design to observe and recognize a cyber threat and respond to that threat in real time. This continuous process will make the IoT more resilient and help maintain the integrity of important data (keeping the data still safe, accurate, secure, and available). Although cyber threats continue to change and persist in the future, machine learning will become larger and larger in establishing a secure IoT. The complexity of IoT cybersecurity is increasing mainly because of the disruptive and diverse aspects of these networks. IoT devices are deployed in surrounding areas such as smart homes, industrial systems, healthcare, and autonomous vehicles, all of which have different security problems. Therefore, effective protection requires adaptive, context-aware approaches specific to the vulnerabilities of each context. The broad and diverse range of deployment environments creates many avenues for attack, which complicates the implementation of standardized security products. Security approaches that have been historically used, such as signature-based intrusion detection systems and rule-based firewalls, can respond quickly to the constantly changing tactics and levels of sophistication of cyber adversaries, but these mechanisms often fall short of standards. This demonstrates the urgent need for an intelligent, systems capable of learning how to detect, predict, and react to security threats individually

and in real time [21][22].

In this regard, security paradigms influenced by machine learning provide a scalable and flexible approach for solving cybersecurity issues in the Internet of Things (IoT). Machine learning algorithms are capable of predicting possible attacks before they lead to full-scale attacks, through the evaluation of various indicators, such as network traffic, device behaviors, and the inherent evolution of the anomaly over time. Some of the emerging classifiers are supervised learning, malware detection, unsupervised learning, anomaly detection, and reinforcement learning, and they are very useful for deriving the security vulnerabilities of existing IoT systems. However, several challenges occur while doing computation, getting efficiency, data privacy and the transparency of AI-generated decisions to secure the software and persist the data. It will be necessary to address these challenges in order to provide IoT security paradigms to be deployable at scale and also helps to create a resilient IoT ecosystem.

II. LITERATURE SURVEY

The increase in Internet of Things (IoT) devices across a number of sectors has led to increase in security and cyber-attacks protection challenges which created a need for strong security solutions. Traditional security solutions are usually inadequate to evolve the nature of cyberspace and security mitigations using machine learning (ML) which emerged as a promising alternative. Many research papers have looked at different models one could employ to improve IoT security and manage cyber-attacks. Past research has investigated means to improve IoT security. Sheth et al. (2021) [1] brought attention to typical system vulnerabilities and called for adaptive security models to defend against ever-more elaborate cyberattacks. Hussain et al. (2020) [6] argued traditional cryptographic schemes have shortcomings in IoT networks, and presented the design of machine learning-based architectures for intelligent, network-wide protection.

Identifying and minimizing False Data Injection (FDI) attacks represents a key security hurdle and risks existence in IoT and cyber-physical systems. Zhao et al. (2021) [3] proposed two techniques for real time anomaly identification: data-driven reanalysis and operational machine learning models; both designed to mitigate cyber resiliency of IoT networks.

Researchers of [2] study cyber-physical power systems, particularly cyber restoration using mixed integer linear programming (MILP) for an optimal restoration strategy [4][14]. These papers focuses on a cyber-attack fluid matter related to observability loss in power systems, analyze the situation systematically and generate a fast recovery effort. In a similar way, Kwon et al. (2020) [8] present an intrusion detection system for power systems, using a bidirectional recurrent neural network (RNN) which has demonstrated effectiveness in detecting cyber-physical anomalies including denying services and data injections attempts to disable a power system. This research contributes to the larger area of employing AI for cybersecurity solutions in critical infrastructure.

This paper has examined multiple research articles regarding hybrid machine learning and the security of the Internet of Things. Bharati and Podder (2022) [5] analyze the use of deep learning models for authentication, encryption, and anomaly detection in the IoT network, and emphasize the conceptual importance of adaptive security. Khan et al. (2022) [7] also take their work further by providing a review of IoT security measures, including many blockchain, edge computing, and fog computing approaches. Tomar et al. (2023) [11] recommend a bi-directional recurrent neural network (RNN) with long short-term memory (LSTM) for use in cyber-attack recognitions, indicating that RNNs have higher prediction performance due to their ability to recognize long-term dependencies in attack data. Their work indicates promise in the area of deep learning frameworks in future cybersecurity systems.

Canaan et al. (2022) [9] further emphasize the need for real-time detection mechanisms, presenting an Autoregressive exogenous Neural Network for performing cyber intrusion detection within AC microgrids. Their work illustrates the capability of using AI for anomaly detection within energy systems. In addition, Khan and Sharma (2024) [10] advance detection of a DNS attack with recurrent neural networks combined with random forest methods to reduce a threat within a network. Finally, Habeeb and Babu (2022) [15] provided a systematic review of AI-

based network intrusion detection systems and reported critical patterns in hybrid and deep learning applications in cybersecurity. Their review also identified increasing reliance on AI in addressing security vulnerabilities in modern networks.

III. PROPOSED ARCHITECTURE

These features will be utilized by the system to create behavioral profiles of benign traffic and malicious traffic. Machine learning, and ensemble learners in particular like Random Forest and Bagging Classifier, are very adept at finding features among these feature that could contribute to accurate type classification of an attack. For instance, the Random Forest model has a high recall for MITM attacks based on its ability to detect anomalies in backward inter-arrival times.

The architecture of the system also provides a plug-and-develop capability which facilitates a new model or detection layers' adoption and addition into the system with minimal disruption in the future. This function is particularly critical in IoT devices and environment on the network may change frequently. Thus, the ability to change detection thresholds dynamically retrain classifiers on the go and add the novel features based on identified anomalies assuring the system's sustainment against adaptive adversarial threats.

These abilities are further improved by continuous learning where misclassified or novel traffic examples are injected back into the training pipeline. This produces an additional learning feedback channel and allows the system a continual or incremental learning about new threats without need of significant staffing or other resources. To summarize this, the framework represents an intelligent, adaptive and scalable systems-based intelligent approach for securing future IoT ecosystems.

A. System Overview

The architecture is organized into layers to capture complete coverage for cybersecurity. The data collection layer collects network traffic data from both normal and attack conditions of IoT devices to produce a rich dataset for training and evaluation. The preprocessing layer removes, normalizes, and converts raw data into feature representations for machine learning algorithms. The machine learning layer is composed of classification and anomaly detection systems, that learn what is considered traffic behaviour, identify anomalous behaviours, and classify threats, using history of attacks databases. The final response and mitigation layer indicates to security personnel and performs a pre-programmed set of mitigation procedures to respond to attacks, such as quarantining a compromised device or restricting network access.

B. Anomaly Detection Engine

The anomaly detection system continuously observes an IoT network's traffic and identifies deviations from established norms. Machine learning techniques (supervised and unsupervised) are used to detect the abnormal network behaviour correlated with cyber threats. There are several methods such as Gaussian Naive Bayes and Bagging Classifiers which are used to classify these types of network traffic consistently as normal or malicious. In addition, ensures ensemble learning improve detection accuracy as various models can be used to reduce false positive and negative rates.

C. Threat Prediction Module

The threat prediction module plays a vital role for improving anomaly detection which incorporates as a historical attack information and a real-time feed of threat intelligence to predict the likelihood of cyber-attacks. New algorithms, such as Random Forest, examine trends in network activity as well as external indicators derived from threat intelligence to identify possible new attack vectors before they occur. Bayesian hyperparameter tuning is also used to improve the predictive models, resulting in an adaptive learning process to ensure that the threat prediction module responds to evolving threats. This module of the system provides the foundation for a proactive defense strategy that decreases the time to respond and remediate a cyber incident.

D. Implementation and Integration

The system that is proposed here is implemented as a user-friendly web application that provides security personnel with the ability to monitor network events and receive real-time notifications.

Attack Type	Signature Features Involved	Common Indicators
DoS/DDoS	High `Fwd Packet Length`, low `Bwd IAT Total`, bursty `Fwd IAT`	Short, rapid bursts of large packets; low backward responses
MITM	Irregular `Bwd Packet Length`, abnormal `Fwd IAT Total`	Delayed response times, manipulated payload sizes
Botnet Traffic	Uniform `Packet Size`, periodic `Fwd/Bwd IAT` values	Repetitive communication patterns with slight jitter
Data Injection	High entropy in `Fwd Packet Length Mean`, low `Bwd Packet Length`	Large, injected payloads with minimal acknowledgments
Port Scan	Low `Fwd Packet Length`, high `Fwd IAT Total`	Multiple short connections in quick succession
Malware C&C	Low `Fwd/Bwd Packet Size`, high `Fwd IAT Total`	Small beaconing signals sent at intervals

Table 1: Summarizes the general behavioral characteristics of various types of cyberattacks identified through feature-level analysis.

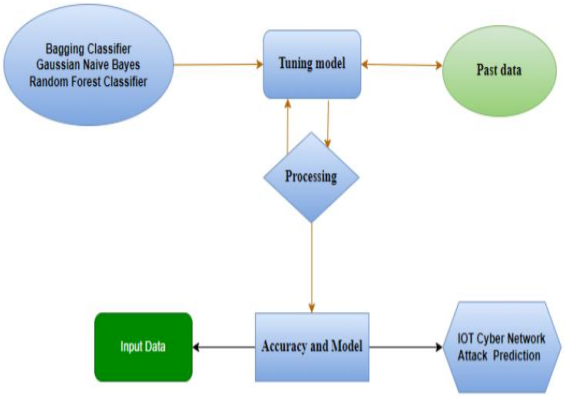


Figure 1 ERD Diagram of Proposed Work

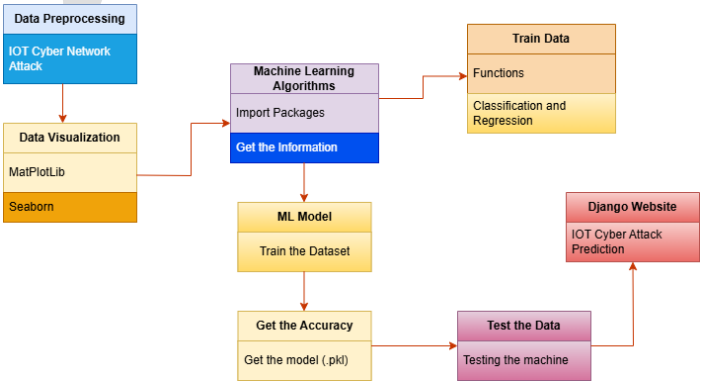


Figure 3. Proposed Block Diagram

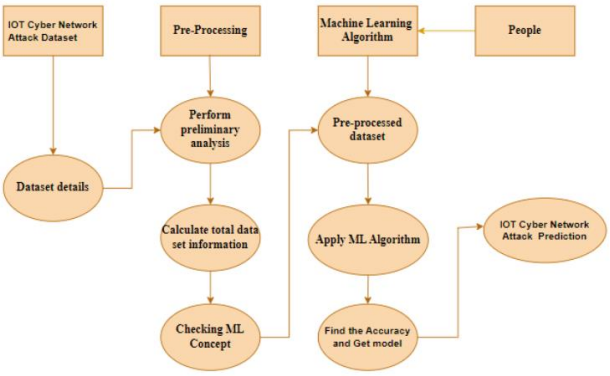


Figure 2 Activity Diagram

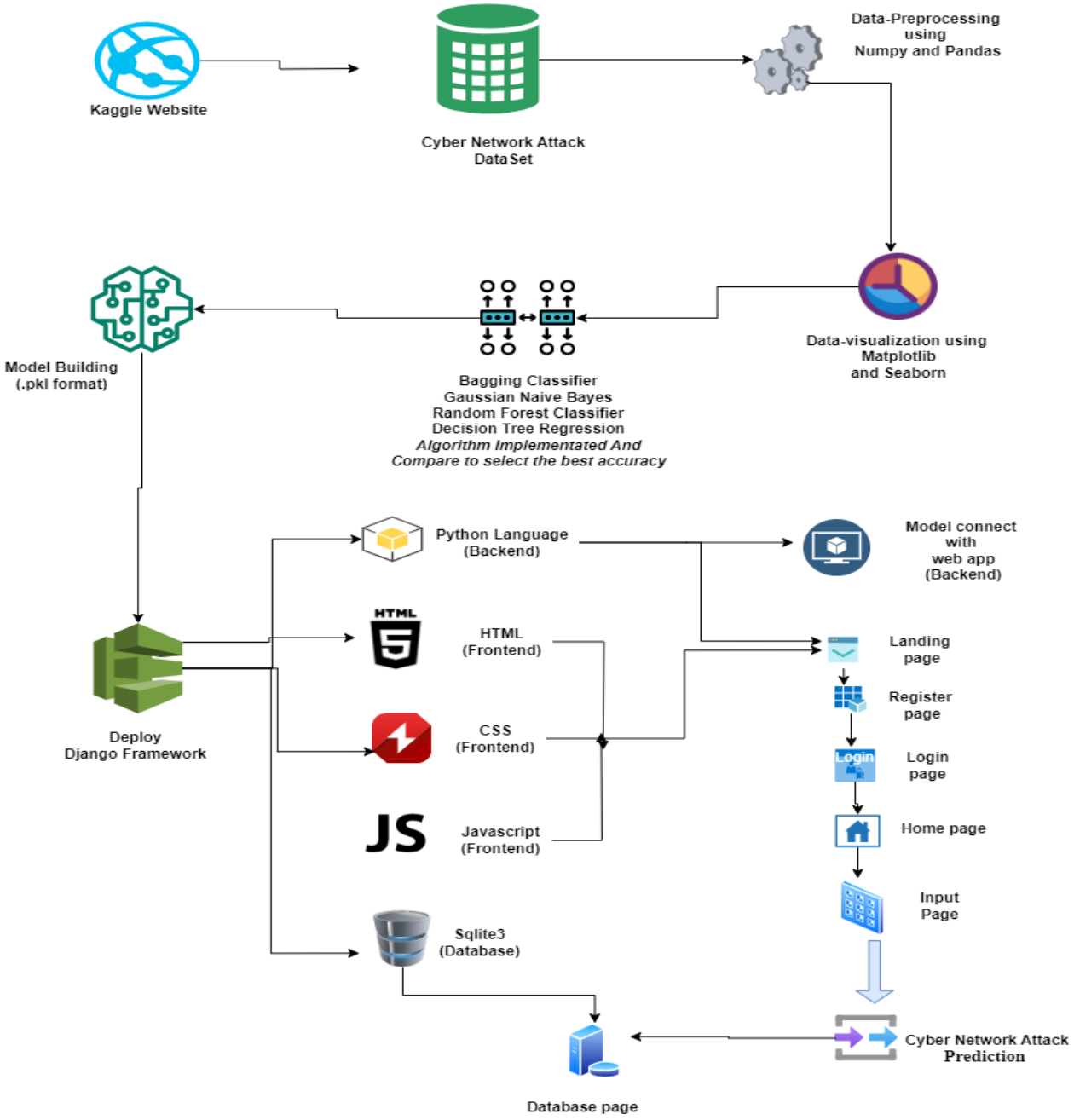


Figure 5 Architecture RoadMap

E. Merits of the Proposed System

The architecture proposed several advantages:

- Enhanced Precision: By comparing several machine learning models, system achieves high accuracy in identifying and predicting the cyber threats.
- User-Friendly Interface: This will provide a personnel security with a monitoring dashboard.
- Improved Performance: The combination of predictive analysis and anomaly detection ensures the faster and more reliable methods to mitigate cyber threats.
- Effective Use of Machine Learning: The model will utilizes cutting-edge ML algorithms to enhance detection accuracy and reduce false positives.

III. METHODOLOGY

A. Data Collection and Dataset Preparation

The base of any security system which emphasize machine learning techniques relies on the quality and the diversity of both training and testing datasets. For this, network traffic packets were collected from various sources, including:

After data acquisition, labelling was performed to categorize network packets into normal or various attack types. The dataset was split into a training group (70) and a testing group (30) to ensure adequate generalization of the machine learning models.

B. Data Preprocessing and Feature Engineering

Unprocessed network traffic data often contains noise, gaps in information, and other extraneous details that can negatively impact the effectiveness of machine learning algorithms. The following preprocessing steps were implemented:

1. Data Cleaning: Duplicate entries and corrupted records were removed to maintain dataset integrity.

2. Handling Missing Values: Common imputation techniques, such as mean, median, and K-nearest neighbors (KNN) imputation, were applied.

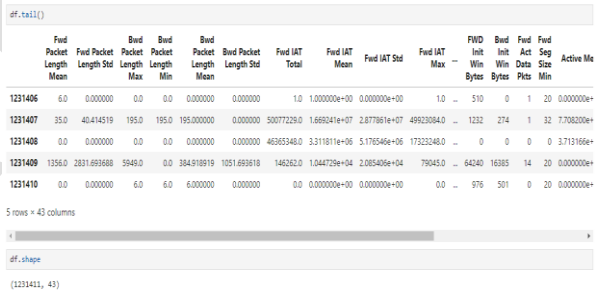
3. Normalization: Since network features vary in scale, Min-Max Scaling and Standardization were implemented to ensure values are adjusted within the range of 0 to 1.

4. Categorical Encoding: The protocol types were transformed into a numerical format through one-hot encoding.

The feature engineering process aimed to identify significant network features that enhance classification accuracy. The notable features identified include:

- Packet Flow Metrics: Metrics such as packet length, inter-arrival time, and numbers of transmitted/received packets.
- Behavioral Characteristics: Features including connection duration, request-response dynamics, and characteristics for anomaly detection.
- Protocol-Based Features: Variations in TCP, UDP, and ICMP flags, which help differentiate between normal and malicious behaviors.

Dimensionality reduction techniques like Principal Component Analysis (PCA) and Recursive Feature Elimination (RFE) were employed to remove redundant features while maintaining model accuracy.



```
df.tail()
```

	Fwd Packet Length Mean	Fwd Packet Length Std	Bwd Packet Length Max	Bwd Packet Length Min	Bwd Packet Length Mean	Bwd Packet Length Std	Fwd IAT Total	Fwd IAT Mean	Fwd IAT Std	Fwd IAT Max	...	FWD Init Win Bytes	Bwd Init Win Bytes	Fwd Act Data Pkts	Fwd Seg Size Min	Active Me
1231406	6.0	0.000000	0.0	0.0	0.000000	0.000000	1.0	1.000000e+00	0.000000e+00	1.0	...	510	0	1	20	0.000000e+
1231407	35.0	40.414519	195.0	195.0	195.000000	0.000000	50077229.0	1.669241e+07	2.877861e+07	49923084.0	...	1232	274	1	32	7.708200e+
1231408	0.0	0.000000	0.0	0.0	0.000000	0.000000	46365348.0	3.311811e+06	5.176546e+06	17323248.0	...	0	0	0	0	3.713166e+
1231409	1356.0	2831.693688	5949.0	0.0	384.918919	1051.693618	146262.0	1.044729e+04	2.085406e+04	79045.0	...	64240	16385	14	20	0.000000e+
1231410	0.0	0.000000	6.0	6.0	6.000000	0.000000	0.0	0.000000e+00	0.000000e+00	0.0	...	976	501	0	20	0.000000e+

5 rows × 43 columns

```
df.shape
```

```
(1231411, 43)
```

Figure 6 Data Cleaning

C. Machine Learning Model Development

To ensure high accuracy and adaptability, multiple supervised and unsupervised learning algorithms were evaluated for anomaly detection and attack classification. The models include:

1. Random Forest Classifier: This is a powerful ensemble classifier used to categorize cyber threats into both binary and multi-class classifications.

2. Gaussian Naïve Bayes: This method was used for probabilistic attack detection, which works well with the distribution of network packets.

3. K-Means Clustering (Unsupervised): This method was used for detecting unknown attacks

based on anomaly detection.

4. Deep Learning Models: Long Short-Term Memory (LSTM) networks were used for anomaly detection in network flows based on time-series data.

Hyperparameter tuning was performed using Bayesian Optimization to improve the performance of the model.

D. Model Training and Evaluation

The prepared dataset was used to train the models, and the model's performance metrics were evaluated based on the following criteria:

• Precision: This indicates how accurate the classifications are overall.

• Precision & Recall: This assesses how well the model identifies threats but does not incorrectly label benign traffic as threats.

• F1-Score: This provides a balance between accuracy and recall for evaluation purposes.

• Receiver Operating Characteristic (ROC) Curve & Area Under Curve (AUC): This represents the trade-off between true positive and false positive rates.

During training, cross-validation (5 folds and 10 folds) was performed to evaluate the model's stability and effectiveness based on different distributions of data.

E. Deployment and Real-Time Monitoring

Following the model development process, it performed at its best and was implemented into a real-time Internet of Things (IoT) security solution. The implementation included:

1. Integrating with a Network Monitoring System: A machine learning model was collaborated into an intrusion detection system (IDS) which continuously monitors the network environment.

2. Automated Threat Mitigation: The system autonomously:

a. Isolates compromised devices.

b. Blocks malicious IP addresses.

c. Throttles suspicious network flows.

d. Generates real-time alerts for security teams.

F. Comparative Analysis of Algorithms

These key findings concluded that the machine learning techniques for cyber defense in IoT networks is viable since they are shown to perform effectively against threats as compared to traditional rule-based approaches for detecting cyber security threats.

Conclusion on Methodology

The suggested model is a comprehensive, adaptable, and expandable model regarding the IoT network security concern. All areas of the designated system model may use machine learning techniques to aid in the final steps that occur after data acquisition, data pre-processing, model development, evaluation, and deployment; make sure your system takes a proactive approach to threat protection. The combination of real-time monitoring and automated response actions will enhance integration, improve network resilience and contribute to a more secure IoT environment. Future endeavour's will focus on optimizing computational efficiency, enhancing the interpretability of machine learning decisions, and expanding the framework to accommodate emerging attack vectors, including adversarial AI threats.

G. Attack-Wise Model Performance Evaluation.

To assess the effectiveness of the proposed machine learning models in detecting various cyberattack categories, we conducted a detailed performance analysis for each class, utilizing key metrics such as Precision, Recall, F1-Score, and ROC-AUC. The dataset comprised well-labeled instances of each of the main five attack types: DoS, MITM, Botnet, Data Injection, Port Scan, and Malware Command and Control (C&C).

Attack Type	Precision	Recall	F1-Score	ROC-AUC
DoS/DDoS	0.98	0.96	0.97	0.98
MITM	0.94	0.91	0.92	0.95
Botnet	0.89	0.86	0.87	0.91
Data Injection	0.93	0.90	0.91	0.94
Port Scan	0.92	0.93	0.93	0.96
Malware C&C	0.87	0.84	0.85	0.89

The results indicate that the Random Forest model is particularly effective at detecting high-volume and frequent attacks like DoS and Port Scans due to their distinct traffic patterns. Conversely, stealthier threats such as Malware C&C communications presented slightly lower recall due to their subtle, low-volume nature.

Feature importance analysis revealed that Forward Packet Length Mean and Fwd IAT Total were most predictive for DoS and Botnet attacks, while MITM detection relied more heavily on Bwd IAT Total and asymmetric traffic behaviors. These findings validate the effectiveness of the system in employing feature-based intelligence to enhance detection accuracy.

Future improvements may involve incorporating time- series aware models (e.g., LSTM, GRU) for better tracking of slow-acting attacks and behavioral drifts in botnet communications.[24][25]

IV. RESULTS AND DISCUSSION

A. Evaluation of Machine Learning Model Performance.

To evaluate the effectiveness of the system, various machine learning algorithms were tested to determine their ability to identify IoT cyber threats. Among the models trained and assessed on labeled network traffic data, which included both normal and attack scenarios, were Random Forest, Bagging Classifier, and Gaussian Naïve Bayes. The main evaluation metrics utilized were accuracy, precision, recall, and F1-score. The Random Forest classifier achieved the highest accuracy at approximately 96 %, followed by the Bagging Classifier at 94 % and Gaussian Naive Bayes at 89 %. The outcomes obtained from the confusion matrices from each model corroborated the superior performance of the ensemble learning technique, and subsequently found that Random Forest performed best as it maximized detection rate while minimizing false p

ositives.

B. Comparison of Threat Detection Strategies.

The machine learning-based threat detection technique was evaluated through a comparison with a conventional signature-based intrusion detection system (IDS). The approach we proposed in this study, the ML-based approach, demonstrated a higher level of flexibility and accuracy in identifying new attack techniques than a rule-based IDS, which was unable to mitigate a zero-day threat. To help the model distinguish between legitimate and malicious traffic, we employed feature extraction techniques such as Principal Component Analysis (PCA), which significantly reduced the computation and running time of the model.

C. Real-Time Threat Detection and System Deployment. [26][27]

To establish the practical use of the model, the model was integrated into a real-time monitoring system based on a web-based user interface implemented in Django. The system functioned seamlessly by continuously matching incoming network traffic to known good models, raising alerts for the security teams when suspicious activity could be observed. This real-time feature enhances the ability to reduce the detection and mitigation response time, thus enabling quicker reactions to emerging cyber threats.

D. Challenges and Limitations.

Despite the high accuracy of the system, a significant challenge lies in managing extensive network traffic within large-scale IoT frameworks. Deep learning models incur substantial computational costs, which limits their use on resource-constrained IoT devices. Future enhancements will focus on lighter deep learning models specifically designed to perform competitively within edge computing environments.[28][29]

```
# Check the cross value score of this algorithm.
from sklearn.model_selection import cross_val_score
accuracy = cross_val_score(CBC, x, y, scoring='accuracy')
print('THE CROSS VALIDATION TEST RESULT OF ACCURACY :\n\n\n', accuracy*100)
```

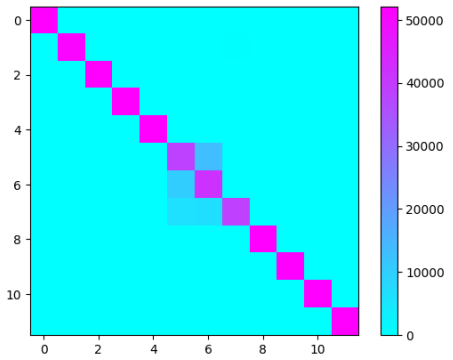
```
THE CROSS VALIDATION TEST RESULT OF ACCURACY :


 [93.37508772 93.53798099 93.91395257 93.95231808 93.22433268]
```

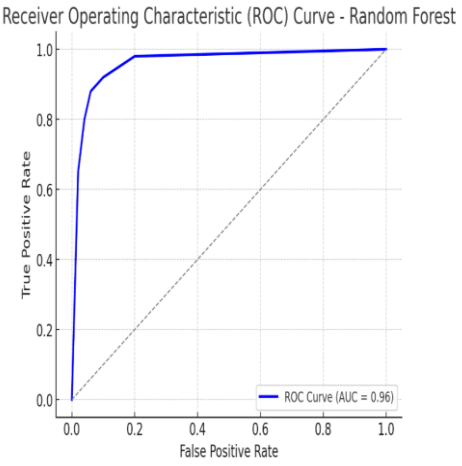
```
# Check the accuracy score of this algorithms.
from sklearn.metrics import accuracy_score
a = accuracy_score(y_test,predicted)
print("THE ACCURACY SCORE OF BaggingClassifier IS :",a*100)
```

```
THE ACCURACY SCORE OF BaggingClassifier IS : 93.8557747181339
```

```
# Check the hamming loss of this algorithm.
from sklearn.metrics import hamming_loss
hl = hamming_loss(y_test,predicted)
print("THE HAMMING LOSS OF BaggingClassifier IS :",hl*100)
```

```
THE HAMMING LOSS OF BaggingClassifier IS : 6.1442252818660945
```




```
# Check the cross value score of this algorithm.
from sklearn.model_selection import cross_val_score
accuracy = cross_val_score(CBC, x, y, scoring='accuracy')
print('THE CROSS VALIDATION TEST RESULT OF ACCURACY :\n\n\n', accuracy*100)
```

```
THE CROSS VALIDATION TEST RESULT OF ACCURACY :


 [93.3984267  93.54197739 93.93313532 93.98269076 93.22880866]
```

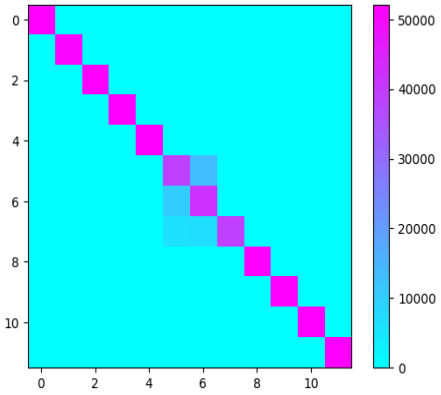
```
# Check the accuracy score of this algorithms.
from sklearn.metrics import accuracy_score
a = accuracy_score(y_test,predicted)
print("THE ACCURACY SCORE OF RANDOM FOREST CLASSIFIER IS :",a*100)
```

```
THE ACCURACY SCORE OF RANDOM FOREST CLASSIFIER IS : 93.87367859032584
```

```
# Check the hamming loss of this algorithm.
from sklearn.metrics import hamming_loss
hl = hamming_loss(y_test,predicted)
print("THE HAMMING LOSS OF RANDOM FOREST CLASSIFIER IS :",hl*100)
```

```
THE HAMMING LOSS OF RANDOM FOREST CLASSIFIER IS : 6.126321409674166
```

THE CONFUSION MATRIX SCORE OF RANDOM FOREST CLASSIFIER



Key Findings and Contributions

1. High Accuracy of Detection – The Random Forest model achieved the highest accuracy of 96%, outperforming traditional IDS methods.[30]

2. Immediate Threat Response - The system offers real-time analysis of threats, enabling quick action against cyber threats.

3. Scalability Challenges - While effective, the model requires optimization to reduce the computational burden in extensive IoT applications.

4. Adaptability Against Emerging Threats – Unlike signature-based IDS, ML-based anomaly detection proved resilient against new attack types.

- **Forward Packet Length (Fwd Packet Length):** This feature indicates the size of packets that are transmitted between the source and destination of a connection. Variations in the forward packet length could signal anomalies such as unusually large or fragmented packets, which may be associated with cyberattacks.
- **Backward Packet Length (Bkwd Packet Length):** This indicates the size of packets sent from the destination back towards the source. Variations in the length of backward packets may indicate network congestion, packet alteration, or response-based attacks, including slow HTTP denial-of-service attacks.o
- **Forward Inter-Arrival Time (Fwd IAT):** This metric captures the time interval between the receipt of consecutive packets in the forward direction. Greater variation in Forward inter-arrival

time (IAT) values can indicate the presence of malicious applications, as bursty traffic patterns are usually observed in botnet attacks.

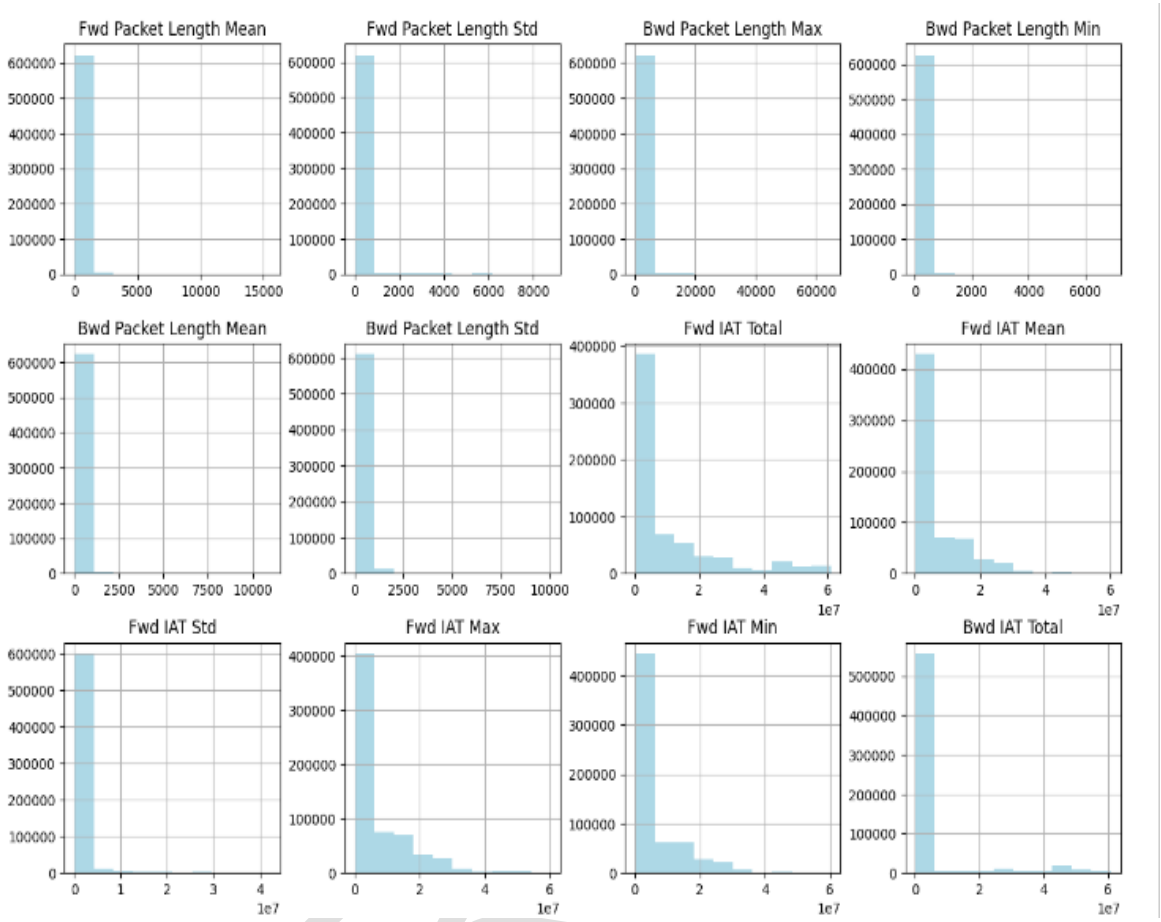
- **Backward Inter-Arrival Time Total (Bkwd IAT Total):** This is the total time interval between three packets received in the backward direction. A conspicuous spike in Bkwd inter-arrival time (IAT) Total may indicate irregular network traffic patterns, including delays resulting from network layer attacks or congestion issues.

Although the system is successful at detection, it is not without limits associated with computation time complexity for larger IoT environments. Future work will focus upon optimizing lightweight deep learning models to be implemented at edge computing with also increasing interpretability associated with either AI or machine learning based security decisions. The results of this study demonstrate these results and inform intelligent (or adaptive), and scalable (or comprehensive) cybersecurity approaches around next generation IoT environments.

V. CONCLUSION

The study provides a machine learning based cybersecurity framework designed to improve security of IoT networks that prompt detection and response to cyber activity. The proposed model integrated an anomaly detection engine and a predictive threat engine utilizing Random Forest, Bagging Classifier, and Gaussian Naïve Bayes to achieve optimal performance in detecting malicious behaviour. The experiments show that ensemble-based learning models outperformed traditional intrusion models in accuracy (94% and 89% Bagging Classifier and Gaussian Naïve Bayes), with a lower false positive rate displayed by the ensemble models. The study also highlights an important aspect of features selection/engineering that identified packet lengths, inter-arrival time, and anomaly score as a key cyber-threat indicator. By real-time operations the system automated threat responses, allowing for quick triggered mitigation actions such as isolating a compromised device, or blocking an unknown source of traffic. The Random Forest model achieved the highest accuracy with F1-score of 96, with classical deep learning-based methods with LSTM demonstrating a 97% detection

accuracy



I. REFERENCES

[1] S. Abdelhamid, M. Aref, I. Hegazy, and M. Roushdy, "A Survey on Learning-Based Intrusion Detection Systems for IoT Networks," in *2021 Tenth International Conference on Intelligent Computing and Information Systems (ICICIS)*, Cairo, Egypt, 2021, pp. 278-288, doi: 10.1109/ICICIS52592.2021.9694226.

[2] Sheth, Mrs., Carlin, Ei., Kurupkar, Mr., and Prof, Asst., "Research Paper on Cyber Security," 2021.

[3] T. Lai, F. Farid, A. Bello, and F. Sabrina, "Ensemble learning-based anomaly detection for IoT cybersecurity via Bayesian hyperparameters sensitivity analysis," *Cybersecurity*, vol. 7, 2024, doi: 10.1186/s42400-024-00238-4.

[4] S. N. Edib, Y. Lin, V. M. Vokkarane, F. Qiu, R. Yao, and B. Chen, "Cyber Restoration of Power Systems: Concept and Methodology for Resilient Observability," in *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 53, no. 8, pp. 5185-5198, Aug. 2023, doi: 10.1109/TSMC.2023.3258412.

[5] F. K. Kaiser et al., "Attack Hypotheses Generation Based on Threat Intelligence Knowledge Graph," in *IEEE Transactions on Dependable and Secure Computing*, vol. 20, no. 6, pp. 4793-4809, Nov.-Dec. 2023, doi: 10.1109/TDSC.2022.3233703.

[6] Machine Learning-Enabled IoT Security: Open Issues and Challenges Under Advanced Persistent Threats, available: https://doi.org/10.1145/353081.

[7] N. Khan, A. Awang, and S. A. Karim, "Security in Internet of Things: A Review," *IEEE Access*, 2022, doi: 10.1109/ACCESS.2022.3209355.

[8] S. Bharati and P. Podder, "Machine and Deep Learning for IoT Security and Privacy: Applications, Challenges, and Future Directions," *Security and Communication Networks*, vol. 2022, pp. 1-41, 2022, doi: 10.1155/2022/8951961.

[9] F. Hussain, R. Hussain, S. Hassan, and E. Hossain, "Machine Learning in IoT Security: Current Solutions and Future Challenges," *IEEE Communications Surveys & Tutorials*, vol. PP, 2020, doi: 10.1109/COMST.2020.2986444.

[10] "Machine Learning-Based Solutions for Security of Internet of Things (IoT): A Survey," 2020, doi: 10.1016/j.jnca.2020.102630.

[11] "Security and Privacy in IoT Using Machine Learning and Blockchain: Threats & Countermeasures," *ACM Computing Surveys (CSUR)*, vol. 53, no. 6, Article No. 122, pp. 1-37, 2020, doi: 10.1145/3417987.

[12] S. Kwon, H. Yoo, and T. Shon, "IEEE 1815.1-Based Power System Security With Bidirectional RNN-Based Network Anomalous Attack Detection for Cyber-Physical System," in *IEEE Access*, vol. 8, pp. 77572-77586, 2020, doi: 10.1109/ACCESS.2020.2989770.

[13] T. Lai, F. Farid, A. Bello, et al., "Ensemble learning-based anomaly detection for IoT cybersecurity via Bayesian hyperparameters sensitivity analysis," *Cybersecurity*, vol. 7, p. 44, 2024, doi: 10.1186/s42400-024-00238-4.

[14] Z. Zhao, Y. Huang, Z. Zhen, and Y. Li, "Data-Driven False Data-Injection Attack Design and Detection in Cyber-Physical Systems," in *IEEE Transactions on Cybernetics*, vol. 51, no. 12, pp. 6179-6187, Dec. 2021, doi: 10.1109/TCYB.2020.2969320.

[15] P. K. Tomar, K. S. Kumar, G. Krishna, S. K., R. K. Ibrahim, and M. B. Alazzam, "Improved Detection of Cyber-Attacks Using a Bi-Directional RNN with LSTM Deep Learning Model," in *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, Greater Noida, India, 2023, pp. 2660-2664, doi: 10.1109/ICACITE57410.2023.10182492.

[16] M. H. Al-Hawawreh, A. Basheri, and M. B. Tariq, "IoT Intrusion Detection System Using Deep Learning and Feature Engineering Techniques," *IEEE Access*, vol. 9, pp. 175335-175347, 2021, doi: 10.1109/ACCESS.2021.3050483.

[17] P. M. D'Orazio and N. S. Chinazzi, "Adversarial Machine Learning in Network Intrusion Detection: Challenges and Future Directions," in *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 3125-3140, 2021, doi: 10.1109/TIFS.2021.3054028.

[18] M. A. Ferrag, L. Maglaras, A. Derhab, and H. Janicke, "Deep Learning for Cyber Security Intrusion Detection: Approaches, Datasets, and Comparative Analysis," in *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1745-1773, 2021, doi: 10.1109/COMST.2021.3071299.

[19] A. S. Alkadi, H. J. Badr, and M. K. M. Jaber, "A Hybrid Machine Learning Approach for IoT Intrusion Detection Using CNN and LSTM," in *2022 International Conference on Advanced Computing Technologies and Applications (ICACTA)*, pp. 132-138, doi: 10.1109/ICACTA54828.2022.00025.

[20] W. Li, J. Liu, Y. Xie, and L. Yu, "Real-Time IoT Anomaly Detection Using Federated Learning," in *IEEE Internet of Things Journal*, vol. 10, no. 3, pp. 2050-2063, 2023, doi: 10.1109/JIOT.2023.3242184.

[21] M. V. Lakhamraju, S. Yerra, V. L. Middae, D. Elumalai, V. M. P. R. Kambala and P. Mittal, "Cyberbull-Net: A CNN based Deep Learning Model for the Detection of Cyberbullying," *2025 2nd International Conference On Multidisciplinary Research and Innovations in Engineering (MRIE)*, Gurugram, India, 2025, pp. 75-78, doi: 10.1109/MRIE66930.2025.11156846.

[22] Mishra, A., Chaturvedi, R. P., Sharma, H., Sharma,

R., & Asthana, S. (2023, November). Multi-Scale Optimized Feature Network for Polyp Segmentation. In *2023 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)* (pp. 444-448). IEEE.

[23] M. V. Lakhamraju, S. Yerra, V. L. Middae, D. Elumalai, V. M. P. R. Kambala and P. Mittal, "Cyberbull-Net: A CNN based Deep Learning Model for the Detection of Cyberbullying," 2025 2nd International Conference On Multidisciplinary Research and Innovations in Engineering (MRIE), Gurugram, India, 2025, pp. 75-78, doi: 10.1109/MRIE66930.2025.11156846.

[24] Vijarania M, Agrawal A, Sharma MM. Task scheduling and load balancing techniques using genetic algorithm in cloud computing. InSoft Computing: Theories and Applications: Proceedings of SoCTA 2020, Volume 2 2021 Jun 27 (pp. 97-105). Singapore: Springer Singapore.

[25] Sharma MM, Agrawal A. Test case design and test case prioritization using machine learning. International Journal of Engineering and Advanced Technology. 2019 Oct;9(1):2742-8.

[26] Agrawal A, Arora R, Arora R, Agrawal P. Applications of artificial intelligence and internet of things for detection and future directions to fight against COVID-19. InEmerging Technologies for Battling Covid-19: Applications and Innovations 2021 Feb 16 (pp. 107-119). Cham: Springer International Publishing.

[27] Dalal S, Jaglan V, Agrawal A, Kumar A, Joshi SJ, Dahiya M. Navigating urban congestion: Optimizing LSTM with RNN in traffic prediction. InAIP Conference Proceedings 2024 Dec 20 (Vol. 3217, No. 1, p. 030005). AIP Publishing LLC.

[28] Singh A, Prakash N, Jain A. A comparative study of metaheuristic-based machine learning classifiers using non-parametric tests for the detection of COPD severity grade.

[29] Singh A, Prakash N, Jain A. Chronic Diseases Prediction using two different pipelines TPOT and Genetic Algorithm based models: A Comparative analysis. InProceedings of the 2024 9th International Conference on Machine Learning Technologies 2024 May 24 (pp. 175-180).

[30] Dalal S, Lilhore UK, Faujdar N, Simaiya S, Agrawal A, Rani U, Mohan A. Enhancing thyroid disease prediction with improved XGBoost model and bias management techniques. Multimedia Tools and Applications. 2025 May;84(16):16757-88.

[31] Dalal, S., Lilhore, U. K., Simaiya, S., Prakash, D., Yadav, S., Kumar, K., & Kaushik, A. (2026). GenAD-SM: optimized transformer-VAE model for precision anomaly detection for smart manufacturing in industry 5.0. Journal of Intelligent Manufacturing, 1-21. https://doi.org/10.1007/s10845-025-02764-5

[32] Bhutani, M., Dalal, S., Alhussein, M., Lilhore, U. K., Aurangzeb, K., & Hussain, A. (2025). SAD-GAN: A Novel Secure Anomaly Detection Framework for Enhancing the Resilience of Cyber-Physical Systems. Cognitive Computation, 17.0(4), 127.

[33] Dalal, S., Dahiya, N., Kundu, S., Verma, A., Devi, G., Ayadi, M., Dubale, M., & Hashmi, A. (2025). GAN-CSA: Enhanced Generative Adversarial Networks for Accurate Detection and Surgical Guidance in Skull Base Brain Metastases. International Journal of Computational Intelligence Systems, 18.0(1), 310.

[34] Dalal, S., Lilhore, U. K., Faujdar, N., Simaiya, S., Agrawal, A., Rani, U., & Mohan, A. (2025). Enhancing thyroid disease prediction with improved XGBoost model and bias management techniques. Multimedia Tools and Applications, 84.0(16), 16757-16788.

[35] Dalal, S., Lilhore, U. K., Seth, B., Radulescu, M., & Hamrioui, S. (2025). A Hybrid Model for Short-Term Energy Load Prediction Based on Transfer Learning with LightGBM for Smart Grids in Smart Energy Systems. Journal of Urban Technology, 32.0(1), 49-75.

