EAI Endorsed Transactions
on Internet of Things

Research Article EAI.EU

Advancing IoT Security with an Innovative Machine Learning Paradigm for Botnet Attack Detection

Punitha P[1], Dinesh Kumar V.K[2] and Lakshmana Kumar R[3,*]

[1]Department of Artificial Intelligence and Data Science, Tagore Institute of Engineering and Technology, Salem, India.
[2]Department of Computer Engineering, NPA Centenary Polytechnic College, The Nilgiris, India.
[3]Department of Artificial Intelligence and Machine Learning, Tagore Institute of Engineering and Technology, Salem, India.

Abstract

INTRODUCTION: In contemporary society, everyday operations are greatly improved by the Internet of Things (IoT), which connects physical devices to provide digital services. IoT technology offers unified services and streamlines activities across various domains, ranging from remote monitoring to sophisticated welfare systems. However, the growing number of IoT devices presents a security concern. Many of these devices are susceptible to exploitation, leading to diverse vulnerabilities.

OBJECTIVES: Resource-constrained IoT devices become prime targets for botnet attacks, manifesting in various forms and penetration methods. Despite numerous research efforts introducing multiple approaches for detecting botnet attacks in IoT, existing methods often fail to achieve satisfactory detection rates.

METHODS: Additionally, these approaches struggle to comprehensively analyze the diverse communication networks within the expansive realm of IoT devices. This study proposes an innovative machine-learning framework for detecting IoT botnet threats to address these limitations.

RESULTS: This conceptual framework exhibits a remarkable capability to identify a spectrum of botnet attacks, showcasing a detection accuracy of 99.5 per cent, significantly surpassing the performance of other prevalent machine-learning approaches.

CONCLUSION: Through this research, we aim to enhance the security paradigm of IoT networks, ensuring robust protection against evolving botnet threats in the dynamic landscape of interconnected devices.

Keywords: Machine Learning, Feature Extraction, Botnets, Internet of Things.

Received on 01 12 2023, accepted on 13 01 2025, published on 03 02 2025

doi: 10.4108/eetiot.4521

1. Introduction

The Internet of Things (IoT) has been a significant development in communication networks within the last ten years, presenting unique challenges in anomaly identification and monitoring across interconnected devices. The intricacies of IoT networks introduce specific issues when discerning anomalies, not all of which necessarily indicate malicious intent. Anomalies can offer valuable insights into traffic behaviour, providing crucial data in various applications, albeit not always causing harm. The scarcity of data with the requisite qualities necessitates a systematic study to parameterize the data, mitigating challenges associated with result validity in anomaly detection.

As the number of Internet-connected devices increases, safeguarding them becomes increasingly complex for organizations, particularly due to the attractiveness of IoT devices to scammers. These devices, vulnerable and prone to security lapses, form an enticing attack surface. IoT security, encompassing techniques to manage internet-connected or network-based groups, faces the challenge of enabling privacy by standard, ensuring the latest

[*] Corresponding author. Email: lakshmanakumar93@gmail.com

operating systems, and utilizing secure equipment during the design phase. Implementing Public Key Infrastructure (PKI) is a crucial strategy to protect client-server associations among diverse network elements. Maintaining a secure IoT network involves enforcing security controls, deactivating port forwarding, avoiding unnecessary exposure of connections, and keeping systems consistently patched and up to date.

Botnet attack detection is more complicated. Most of the attacker's target IoT networks have rapid development and less security concern [1]. Botnet is a combination of the term robot and networking. Bots are machines that conduct malicious activities remotely, such as PCs, laptops, smartphones, servers, static routing, etc. A botnet is a collection of compromised devices that hackers with malicious intent can program to perform coordinated actions. Botnets are critical Internet safety risks responsible for the bulk of spamming, impersonation, and hijacking with several DDoS attacks [2]. In the instance of a functioning botnet, for example, once a sufficient number of devices have been compromised, a bot herder can use the remote control to launch cyberattacks or other harmful objectives. Because a bot herder can disperse attack duties over the Internet, botnet-based assaults are particularly dangerous and expensive to fight off due to the massive aggregate bandwidth and many assault sources [3]. As the Internet expands, the prevalence of botnets in cyber-attacks poses serious threats to network services and users. Anomalies in the open connectivity of IoT networks hinder traditional network anomaly detection methods from effectively identifying botnets. Various IoT devices, such as camcorders, routers, smart sensors, and home appliances, fall victim to botnet infections due to the absence of essential security mechanisms and shared credentials [4].

Anomaly detection aims to identify patterns whose behaviour differs from regular nodes. Intrusion detection systems [5], fraud detection, and data leakage are all potential sources of abnormalities. Application program interface (API) protection, public key infrastructure (PKI) authorization, and cybersecurity are just a few small strategies IT managers can employ to tackle the growing cybersecurity threats and cyberterrorism stemming from insecure IoT devices. Internet Relay Chat (IRC) is a common way for botnets to function, and one way to identify IRC-based botnets is to monitor TCP port 6667, which is the default port for IRC communication. A network of hacked devices, such as computers hacked with bot software, that is under the control of hackers is known as a botnet. More hackers may rent out the bot network to utilize it for [DDoS] attacks [6] and spam distribution.

Detecting botnet attacks adds another layer of complexity, with attackers increasingly targeting IoT networks characterized by rapid development and limited security concerns. A botnet, a fusion of "robot" and "networking," comprises machines conducting malicious activities remotely, such as PCs, laptops, smartphones, servers, and static routing. These networks pose critical internet safety risks, responsible for many issues, including spamming, impersonation, and various (DDoS) attacks [7]. The inherent danger lies in the ability of a bot herder to disperse attack duties over the Internet, making botnet-based assaults particularly dangerous and expensive to counter.

As the Internet expands, the prevalence of botnets in cyber-attacks poses serious threats to network services and users. Anomalies in the open connectivity of IoT networks hinder traditional network anomaly detection methods from effectively identifying botnets. Various IoT devices, such as camcorders, routers, smart sensors, and home appliances, fall victim to botnet infections due to the absence of essential security mechanisms and shared credentials.

Anomaly detection, a pivotal aspect in addressing cybersecurity threats, aims to identify patterns deviating from the behaviour of regular nodes. Intrusion detection systems, fraud detection, and data leakage are potential sources of abnormalities. Strategies like Application Programming Interface (API) protection, Public Key Infrastructure (PKI) authorization, and cybersecurity represent small yet crucial tools for IT managers combating growing threats and cyberterrorism stemming from insecure IoT devices. Botnets, often operated via Internet Relay Chat (IRC), can be detected by monitoring TCP port 6667, the default port for IRC traffic.

The proposed method introduces an innovative machine learning framework for Internet of Things (IoT) environments to identify botnet attacks, distinguished by its innovative feature engineering and adaptability to resource-constrained devices. A key aspect of this framework is its systematic search-based technique for selecting pertinent characteristics from network traffic data, improving the model's capacity to distinguish between normal and malicious activities while reducing computational overhead. Moreover, the framework occupies a hybrid algorithm that combines statistical learning with advanced perception strategies, allowing for a nuanced analysis of network behaviour and improved sensitivity to subtle anomalies indicative of botnet activity. Integrating a core entropy metric refines detection by approximating similarity measures between traffic patterns. Furthermore, the framework is designed for scalability and adaptability, enabling it to evolve with the changing landscape of IoT threats by continuously updating its feature set and detection algorithms. Overall, the method's advanced feature engineering, hybrid analytical capabilities, and high detection accuracy collectively enhance botnet identification in the Internet of Things settings, tackling the particular difficulties brought on by diverse and dynamic networks.

The paper's following sections are organized as follows: Background information is given in Section II, which also highlights the limitations and current state of botnet detection systems as well as the reasons behind the difficulties. Section III details the system integration of our proposed model, while Section IV outlines the assessment results in identifying botnet assaults. Finally, Section V offers a concise overview of our study, emphasizing contributions and proposing avenues for future research.

2. Literature Review

Much research has been conducted in IoT networks to detect botnet attacks. The botnet detection system is designed to analyze and monitor network behaviour and flow. Many approaches and techniques have been developed, which can be explored in the literature review section.

Machine learning models are widely used for botnet attack detection. [8] proposed machine learning model for botnet detection based on graphs. This model is an enhancement of the flow flow-based machine learning model. Flow-based features have a huge computational overhead and do not adequately grasp network communication patterns, one of their main drawbacks. Furthermore, one of the primary flaws of flow-based machine learning algorithms for detecting botnets is that they ignore the dynamic topological information of communications infrastructure. The flow-based machine learning model had heavy overhead and could not recognize the communication patterns. The graph-based algorithm was used to overcome the flow-based ML model's limitations. The graph-based ML model used five filter-based feature evaluations from different theories: consistency, correlation, and information. Even though the proposed model achieves high accuracy and reduces complexity in botnet detection, the model cannot detect the botnets in structural network data.

[9] proposed an Ensemble Intrusion Detection Technique to mitigate botnet attacks in IoT. The proposed method utilized DNS, HTTP, and MQTT protocols to deny botnet attacks with statistical flow features. The AdaBoost ensemble learning method is a combination of three machine learning algorithms: Decision Tree (DT) [10], Naive Bayes (NB) [11] and Artificial Neural Network (ANN) [12]. The Ensemble Intrusion Detection Technique achieves a lower false alarm rate and a higher detection accuracy by differentiating malicious and legitimate activity based on the integrity and association value criteria. A decision tree directly derived from and customized to the current intrusion detection signatures is made using a classification technique. It is possible to swiftly discover all triggering rules with a small number of iterations by employing decision trees to detect intruders. The Naive Bayes classifier will be used to detect threats such as data collecting, denial of service, and the root and distant cause of the assault. The Naive Bayes classifier operates under the assumption of strong independence, meaning that the probability of one characteristic does not influence the likelihood of another. However, the proposed detection technique detects the botnets effectively with the potential features. However, the ensemble technique fails to differentiate between legitimate and malicious attacks. Logistic Regression is an initiative to use quantitative active learning to establish a baseline for an anomaly-based intrusion detection system to improve network security and decrease human participation in botnet detection. Using a lightweight logistic regression model, we intensive on defining the ideal structures for noticing botnet activity inside network activity. Bro, a popular network monitoring platform that gives aggregate data about packets sent between a source and destination over a specified time interval, processes the internet traffic. These metrics are inputs to a logistic regression model that classifies malicious and benign communication.

[13] developed an efficient feature engineering and machine learning model for detecting botnet attacks in IoT. Smart digital services vastly use IoT systems, mainly targeted by malicious attacks [14]. A search-based feature engineering technique was proposed with K-Medoid sampling to avoid such vulnerabilities and protect the IoT network from botnet attacks. Because k-Medoids clustering techniques [15] replicate the real-world scenario of continuous data, the enhanced strategy would more precisely group the entire data into appropriate clusters over K-Means, resulting in an improved categorization. The k-Medoids algorithm, on the other hand, tries to eliminate the squared error, which would be the separation among elements in a cluster and a position chosen as the centroid.

[16] presents an Energy-Efficient packet transmission framework for security in Mobile Adhoc Networks, employing a Multi-Relational Associative Rule Trained Artificial Neural Network. The system employs a multi-relational computation of trust values alongside hashing, clustering, routing, encryption, and prioritized transmission. The experimental results of the model indicate greater effectiveness than traditional methods, showcasing its ability to enhance detection rates, lower false positive rates, and increase energy efficiency in packet transmission. [17] Cognitive Radio Ad Hoc Networks (CRAHNs) are vital in addressing spectrum shortages and rising traffic demands while ensuring high-quality service. Distributing resources is difficult because of user conflicts, traffic congestion, and elevated data transmission error rates. This document presents a reinforcement learning-based ANH model configured by policy to tackle these limitations. The model sets up and clusters nodes with the Link Reliability K-Means algorithm identifies spectrum with CBD and PRW-RBM and employs the Weibull Distribution-based Blue Monkey Optimization method for resource allocation. [18] Cloud computing is crucial for IoT devices, enabling data storage and access. It provides services such as IaaS, PaaS, and SaaS. Nonetheless, it may be susceptible to security threats such as Interior Keywords Guessing Attacks (IKGA). To tackle this, techniques such as Certificateless Fleshed Public Key Authenticated Encryption of Keyword Search (CL-HPAEKS), Modified Elliptic Curve Cryptography (MECC), and Mutation Centered Flower Pollinations Algorithm (CM-FPA) can be employed. The suggested system attains 96% security

and requires less time for implementation. [19] presents a novel anomaly application detection system employing Federated Learning and a Hyperbolic Tangent Radial-Deep Belief Network (FL-HTR-DBN). The system employs the Hadoop System for training, retrieves log files, transforms them into vector forms, and labels them utilizing the SKLD-SED K Means algorithm. The FL-HTR-DBN model is trained with these features and identified anomalies are hashed and stored safely. The system exceeds current precision, recall, accuracy, F-measure, sensitivity, and specificity techniques.

Sandbox environment and collection behaviour are essential in IoT systems [6]. However, the previous sandbox tools could not achieve satisfactory results in identifying botnet attacks. To overcome the limitations of sandboxes [7], proposed V-Sandbox for dynamic analysis of IoT botnet attacks. The V-sandbox has an ideal environment for detecting all types of botnet families. However, the V-sandbox tool is efficient for the dynamic analysis of resource-constrained IoT devices. Table 1 represents the comparison of existing methods.

Table 1. Comparison of Existing Methods

References	**Methodology**	**Results**	**Limitations**
[8]	The paper proposed a graph-based machine learning model for botnet detection, enhancing flow-based models.	The model achieved high accuracy and reduced complexity in botnet detection.	The model could not detect botnets in structural network data; flow-based features had high computational overhead.
[9]	The paper developed an Ensemble Intrusion Detection Technique using DNS, HTTP, and MQTT protocols with statistical flow features.	Mitigated botnet attacks effectively using a combination of three machine learning algorithms.	Limited to specific protocols; may not generalize well across all types of IoT communications.
[2]	The paper developed a feature engineering and machine learning model for detecting IoT-botnet cyber-attacks.	The paper's results significantly improved detection accuracy and reduced false positives.	It may require extensive feature selection and tuning; performance can vary with different datasets.
[6]	The study introduced FlowGuard, an intelligent edge defence mechanism against IoT DDoS attacks.	The results demonstrated effectiveness in mitigating DDoS attacks with low latency.	The study focused primarily on DDoS attacks and may not address other botnet threats.
[11]	The study analyzed the Mirai botnet to understand its behaviour and propagation methods.	The study provided insights into the operational mechanisms of the Mirai botnet, aiding in detection strategies.	This analysis is limited to a specific botnet; findings may not apply to other botnets.
[12]	The study employed hybrid deep learning techniques for botnet attack detection in IoT networks.	The study achieved high detection rates with improved accuracy over traditional methods.	Deep learning models can be resource-intensive and unsuitable for all IoT devices.
[7]	The paper proposed a timely detection and mitigation strategy for stealthy DDoS attacks via IoT networks.	Enhanced detection capabilities for stealthy attacks, improving overall network security.	It may require continuous monitoring and updates to remain effective against evolving threats.

[10]	The study conducted comprehensive research on IoT security, focusing on various attack vectors and defences.	The study identified key vulnerabilities and proposed a framework for improving IoT security.	The broad scope may lead to less focus on specific attack types and implementation challenges in real-world scenarios.

This section reviews existing methods for detecting botnet attacks in IoT systems, including graph-based machine learning, ensemble models, and sandbox tools. While these methods have improved accuracy and reduced complexity, they still face challenges in analyzing dynamic network structures and differentiating legitimate from malicious activity.

3. Proposed Methodology

The proposed framework introduces a machine-learning approach focused on feature engineering and hybrid algorithms. It handles large IoT datasets effectively, improves detection accuracy, and works efficiently on lightweight IoT devices. This method addresses the limitations of traditional approaches by providing real-time anomaly detection and better security for IoT networks.

3.1. Problem Statement

In digital communication, ensuring robust security and fostering trustworthy communication among IoT devices proves to be a complex challenge. IoT devices are extensively employed in diverse sectors, including smart homes, healthcare, and various industrial domains. The compromise of IoT devices poses a significant threat, as they can be exploited for orchestrating botnet attacks within IoT systems. Traditional security measures fail to detect the escalating frequency of botnet attacks in this dynamic environment. Addressing these emerging security and privacy concerns becomes imperative in the context of IoT systems. To understand the intelligence and detect botnet attacks, a novel machine learning-based framework is proposed that recommends feature engineering techniques for improved detection accuracy.

3.2. Proposed Method

Smart applications have rapidly changed with the development of IoT technology. The IoT technology has no built-in security mechanism, and many devices are connected to the IoT systems. It becomes very easy for attackers to attack the IoT systems. A novel machine learning framework was proposed to detect botnet attacks to address the emergency of security issues in IoT systems. With the motivation of literature review approaches, this research proposed a feature engineering-based machine learning framework. These approaches have computing capabilities and costs, especially relative to statistical learning methods, and therefore can be used for lightweight IoT devices. The network data is first preprocessed to select important network flow characteristics. The information is then analyzed and handed to the following stage, which uses a hybrid algorithm and a perception strategy to estimate probability density. A core entropy metric is utilized to approximate similarity measures between the traffic in the next stage. Filtering raw network packets and extracting meaningful information that can aid in efficiently identifying known attacks requires network extraction characteristics. These features still require additional filtration to remove any chaotic or unnecessary data that degrades anomaly detection results regarding attack detection and processing time. Since the most critical features are implemented, it is possible to implement a powerful anomaly detection system at IoT gateways and identify new botnet occurrences in real time. Variations of botnet activities are included in the feature set since they show the most appropriate representation of their behaviours.

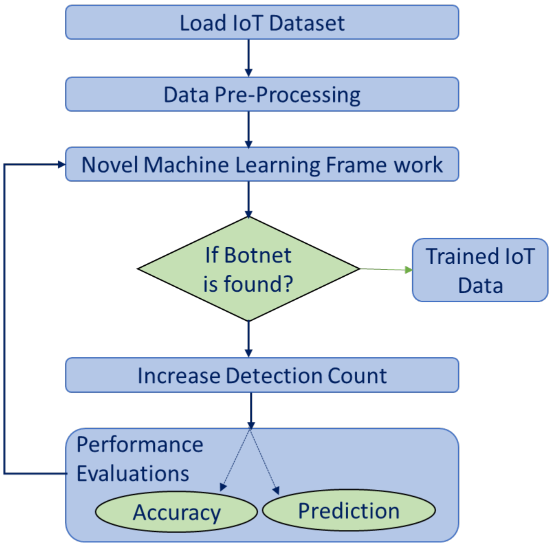

Figure 1. Architecture of Novel Machine Learning Framework

The proposed framework for detecting IoT botnet attacks shown in Fig 1, aims to reinforce the security of IoT networks through a structured and comprehensive approach. The framework integrates key processes, including data collection, preprocessing, feature extraction, and advanced machine learning algorithms, to effectively identify and mitigate malicious activities. The process begins with collecting data from various IoT devices within the network. This data encompasses network traffic logs, device status reports, and communication patterns. The objective is to construct a detailed dataset that captures the normal behaviour of the IoT ecosystem. This baseline understanding of standard operations is essential for identifying anomalies that may signal botnet activities. A diverse and representative dataset ensures a robust foundation for analysis. Once the data is collected, it undergoes a meticulous preprocessing phase to enhance its quality and reliability. This step involves several key actions: cleaning the data by removing null or empty values to maintain the integrity of the dataset; discarding incomplete or erroneous records to ensure consistency; and reducing noise by filtering out irrelevant or chaotic data, which could undermine the performance of detection algorithms.

Preprocessing improves data quality and lays the groundwork for more accurate and efficient analysis by ensuring that only relevant and reliable information is utilized. The next phase focuses on extracting meaningful features from the cleaned network data. These features are designed to encapsulate the distinguishing characteristics of both normal and malicious traffic. A core entropy metric assesses similarity measures among various traffic patterns, facilitating the differentiation between benign and malicious activities. Feature extraction serves two crucial purposes: by selecting the most informative aspects of the data, the framework reduces computational complexity, and by retaining only critical features, it enhances the efficiency and precision of subsequent detection algorithms. The extracted features are fed into a hybrid machine-learning model designed to detect botnet attacks effectively. The model leverages a combination of supervised and unsupervised learning techniques: supervised learning algorithms, trained on labelled datasets, classify traffic as either normal or malicious, leveraging historical data to identify known attack patterns, while unsupervised learning algorithms excel in identifying anomalies without prior labelling, making them valuable for detecting novel or evolving botnet behaviours.

This hybrid approach ensures adaptability to new threats and improves detection rates over time by learning from historical and real-time data. The final component of the framework focuses on real-time detection and response. Machine learning algorithms continuously analyze incoming data, and once a botnet attack is identified, immediate mitigation strategies are deployed. Timely detection and response are critical to minimizing the impact of botnet attacks, ensuring the framework functions proactively and reactively. This framework integrates advanced data processing techniques with robust machine learning capabilities to deliver a reliable solution for detecting IoT botnet attacks. The architecture achieves high detection accuracy and adaptability to new threats by systematically collecting, preprocessing, and analyzing data. This approach enhances the overall security of IoT systems, contributing to developing resilient cybersecurity measures in the rapidly evolving landscape of interconnected devices.

```
Algorithm Name - IoT Botnet Detection Algorithm
Input - Trained IoT Dataset TrID, Test IoT Dataset TsID
Output - Botnet Attack Detection
Start
Load Dataset
Apply Data Preprocess
        {
                Remove Null/Empty Value Records
                Remove N/A Records
        }
Load Preprocess Dataset
IoT Test Dataset $TrID = \sum_{i=1}^{n} RiVi$
Load IoT Train Dataset $TsID = \sum_{j=1}^{n} RjVj$
Apply Classification
        IF( RiVi == RjVj) {
                        Botnet Detected
                        }
                else
                        {
                        Benign
                        }
        End
```

4. Result Analysis

This section evaluates the novel machine learning framework for detecting botnet attacks in IoT systems. The performance assessment involves a comparative analysis with established machine learning algorithms, namely J48, Naive Bayes (NB), Artificial Neural Network (ANN), and Logistic Regression (LR). The proposed framework's classifier spans various machine learning families within supervised classification. Based on machine learning, the trained dataset is then applied within the Novel Machine Learning Framework (NMLF) to detect botnets in the IoT system dataset.

4.1. Experimental Setup

4.1.1. Environment

The experiment of the proposed framework followed different stages such as data analysis, preprocessing, feature selection, preparation of trained model and classification. The best configuration Ai 5 processor and 500 GB storage capacity and 8 GB processing speed super node is used to implement the proposed framework. The proposed framework is implemented using Python programming. The experiment utilized many libraries, such as Pundas and NumPy. Also used graph tools for analysis and visualization of performance

4.1.2. Dataset

The proposed framework for botnet detection is designed to tackle the challenges of securing IoT networks from malicious attacks. Botnets, networks comprising compromised IoT devices, can be exploited by attackers to launch various cyberattacks, such as Distributed Denial of Service (DDoS) attacks and spam campaigns. IoT devices, including smart sensors, routers, and home appliances, are often vulnerable due to their limited built-in security features and the widespread use of shared credentials. This experiment employed real-world botnet traffic data from the N-BaIoT dataset. The framework integrates data preprocessing techniques to clean the dataset by eliminating null values and irrelevant records. Once cleaned, the data undergoes analysis to extract significant features, which are then utilized to train a machine-learning model. The proposed model employs feature engineering to enhance detection accuracy and efficiency. The framework is designed to detect various types of botnets by leveraging a large dataset of IoT device traffic. By comparing test data against a pre-trained dataset, the framework determines whether the observed behaviour indicates a botnet attack or normal activity. The detection process utilizes a hybrid algorithm that combines statistical learning methods with a perception strategy for anomaly detection, facilitating real-time identification of botnet activities.

4.2. Comparative Analysis

Performance is evaluated by comparing different performance metrics provided in the next subsection. Performance metrics such as detection rate and prediction are used to assess the efficiency of the proposed framework.

4.2.1. Accuracy Performance

As shown in Table 2, the results revealed that the proposed framework has significant performance improvement over existing ML algorithms regarding detection rate.

Table 2. Performance of Accuracy

S No	Algorithm Name	Accuracy (%)
1	J48	98.06%
2	Naive Bayes	98.06%
3	Logical Regression	97.22%
4	Artificial Neural Network	96.56%
5	Novel ML Framework	99.34%

When assessing the effectiveness of machine learning models, the accuracy metric is essential because it shows the ratio of accurate predictions to total predictions. Accuracy in IoT botnet detection refers to a framework's capacity to accurately distinguish between malicious and benign activity in IoT network traffic. J48, Naive Bayes (NB), Logistic Regression (LR), Artificial Neural Networks (ANN), and the suggested Novel Machine Learning Framework (NMLF) are among the machine learning algorithms that are compared in Table 1. Each algorithm was tested on a standardized dataset for botnet detection, with accuracy values reflecting their effectiveness. J48 and NB achieved an accuracy of 98.06%, demonstrating their competence in handling structured data but with limitations in adapting to unstructured or dynamic patterns. Logistic Regression showed a slightly lower accuracy of 97.22%, indicating challenges in handling complex, multi-dimensional feature spaces. ANN exhibited an accuracy of 96.56%, which, while robust, may suggest sensitivity to overfitting or insufficient feature engineering. However, the Novel Machine Learning Framework significantly outperformed the existing algorithms with an accuracy of 99.34%. This improvement can be attributed to its advanced feature engineering, entropy-based similarity metrics, and real-time detection capabilities designed for IoT environments. The NMLF demonstrates its ability to process heterogeneous IoT data effectively by employing feature extraction and filtration of network flow characteristics. This hybrid algorithmic approach combines statistical and machine learning methods, and entropy-based metrics for anomaly detection. This enhanced accuracy highlights the framework's suitability for real-time applications, ensuring robust protection against botnet attacks in dynamic IoT networks.

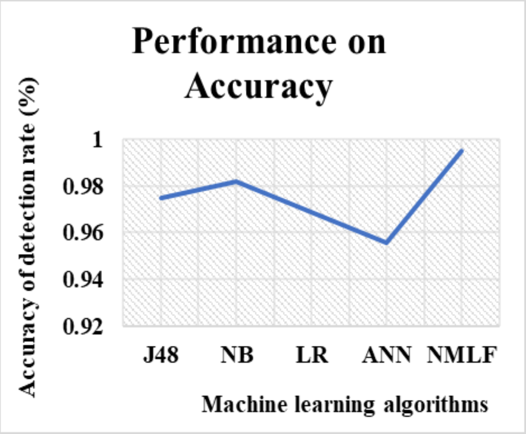

Figure 2. Accuracy Performance

Figure 2 compares the proposed novel machine learning framework's detection accuracy to that of existing machine learning algorithms such as J48, Naive Bayes (NB), Logistic Regression (LR), and Artificial Neural Network (ANN). The graph's horizontal axis represents the various algorithms, while the vertical axis shows the accuracy percentages. The results demonstrate a significant performance improvement for the proposed framework. Specifically, the proposed framework achieves an accuracy of 99.34%, which is higher than the detection rates of the existing algorithms: J48 (98.06%), NB (98.06%), LR (97.22%), and ANN (96.56%). This improvement underscores the framework's enhanced ability to detect botnet attacks effectively within IoT networks. Such superior accuracy highlights the robustness and efficiency of the framework in addressing security challenges compared to traditional machine learning methods [11][12][1]. By leveraging its novel design, the framework outperforms its counterparts, showcasing its potential for practical deployment in securing IoT systems against botnet threats.

4.2.2. Prediction Performance

Table 3. Performance on Prediction Rate

S No	Algorithm Name	Prediction Rate
1	J48	96.13%
2	Naive Bayes	96.78%
3	Logical Regression	97.93%
4	Artificial Neural Network	98.78%
5	Novel ML Framework	99.12%

Table 3 demonstrates a clear improvement in prediction rates when using the proposed framework, Novel Machine Learning Framework (NMLF), compared to other traditional machine learning algorithms. Among the existing algorithms, J48 achieved a prediction rate of 96.13%, showcasing a strong performance but falling short of higher accuracy thresholds. Naive Bayes (NB) achieved a minor enhancement over J48, attaining a prediction rate of 96.78%. Logistic Regression (LR) exhibited a significant advancement with a prediction rate of 97.93%, underscoring its effectiveness in predictive assignments. The Artificial Neural Network (ANN) exceeded the performance of other algorithms, reaching a prediction accuracy of 98.78%. Nonetheless, the Novel Machine Learning Framework (NMLF) accomplished a prediction rate of 99.34%, demonstrating a noteworthy improvement compared to the top-performing current algorithm, ANN. This improvement reflects the superiority of the proposed framework in handling complex IoT data patterns and identifying botnet attacks with higher precision. The incremental increase in prediction rates from J48 to ANN indicates the limitations of traditional models in effectively analyzing IoT datasets. The substantial leap to 99.34% with NMLF highlights the effectiveness of its advanced feature extraction and hybrid algorithms in capturing nuanced data characteristics and reducing false predictions. This performance boost suggests that the proposed framework addresses the gaps left by conventional machine learning methods, particularly in high-stakes applications like botnet detection in IoT systems. The proposed framework achieves a higher prediction rate and underscores the importance of tailored machine-learning solutions for improving cybersecurity in IoT environments. This significant improvement validates the framework's potential for real-world applications where predictive accuracy is critical.

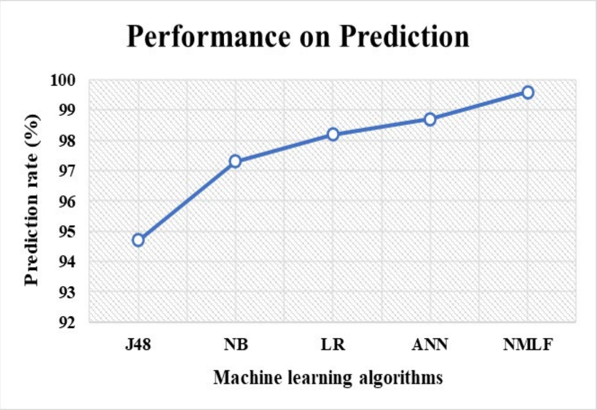

Figure 3. Prediction Performance

The suggested framework's performance in terms of detection rate accuracy is contrasted with that of the current machine-learning techniques, as illustrated in Figure 3. The vertical axis shows the prediction rate, and the horizontal axis lists the names of the algorithms [8][9][10]. The outcomes demonstrated that the suggested structure performed better.

5. Conclusion

The Internet of Things (IoT) makes revolutionary changes in communication technology. The usage of IoT is increasing vastly. The innovative IoT technology enables device interaction and acquires various services. The usage of IoT is gradually increasing in all aspects of information technology. The IoT plays a vital in developing smart applications for multiple industries. Malicious assaults were among the most serious problems that IoT systems confront. Data is collected in real time by IoT-enabled mobile devices. However, most of these do not comply with data networking protocols. The experimental results underscore the efficacy of our approach, with the proposed framework achieving a detection accuracy of 99.34% and a prediction rate of 99.12%. These metrics significantly surpass those of existing algorithms, including J48, Naive Bayes, Logistic Regression, and Artificial Neural Networks. The robust performance is attributed to the optimized feature extraction process and the hybrid algorithm's ability to capture nuanced variations in botnet behaviours. There is a lot of ambiguity when it comes to malevolence and beginning. This research proposed a novel machine learning framework to address the limitation of the system's present state. The proposed framework improved the botnet detection rate significantly in the IoT system.

References

[1] Akmandor AO, Hongxu YIN, Jha NK. Smart, secure, yet energy-efficient, Internet-of-Things sensors. IEEE Trans Multi-Scale Comput Syst. 2018;4(4):914-930.

[2] Panda M, Abd Allah AM, Hassanien AE. Developing an efficient feature engineering and machine learning model for detecting IoT-botnet cyber attacks. IEEE Access. 2021;9:91038-91052.

[3] Mosenia A, Jha NK. A comprehensive study of internet-of-things security. IEEE Trans Emerg Top Comput. 2016;5(4):586-602.

[4] Yin L, Luo X, Zhu C, Wang L, Xu Z, Lu H. ConnSpoiler: Disrupting C&C communication of IoT-based botnet through fast detection of anomalous domain queries. IEEE Trans Ind Inform. 2019;16(2):1373-1384.

[5] Sarker IH, Abushark YB, Alsolami F, Khan AI. Intrudtree: A machine learning-based cyber security intrusion detection model. Symmetry. 2020;12(5):754.

[6] Jia Y. Flowguard: An intelligent edge defense mechanism against IoT DDoS attacks. IEEE Internet Things J. 2020;7(10):9552-9562.

[7] Doshi K, Yilmaz Y, Uludag S. Timely detection and mitigation of stealthy DDoS attacks via IoT networks. IEEE Trans Depend Secure Comput. 2021.

[8] Alharbi A, Alsubhi K. Botnet detection approach using graph-based machine learning. IEEE Access. 2021;9:99166-99180.

[9] Moustafa N, Turnbull B, Choo KKR. An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of Internet of Things. IEEE Internet Things J. 2018;6(3):4815-4830.

[10] Le HV, Ngo QD. V-Sandbox for dynamic analysis IoT botnet. IEEE Access. 2020;8:145768-145786.

[11] Antonakakis M, April T, Bailey M, Bernhard M, Bursztein E, Cochran J, et al. Understanding the Mirai botnet. 26th USENIX Security Symp (USENIX Security 17). 2017:1093-1110.

[12] Popoola SI, Adebisi B, Hammoudeh M, Gui G, Gacanin H. Hybrid deep learning for botnet attack detection in the Internet-of-Things networks. IEEE Internet Things J. 2020;8(6):4944-4956.

[13] Dange S, Chatterjee M. IoT botnet: The largest threat to the IoT network. Data Commun Netw. Springer, Singapore; 2020:137-157.

[14] Wang TS, Lin HT, Cheng WT, Chen CY. DBod: Clustering and detecting DGA-based botnets using DNS traffic analysis. Comput Secur. 2017;64:1-15.

[15] Termanini RD. The Nano Age of Digital Immunity Infrastructure Fundamentals and Applications: The Intelligent Cyber Shield for Smart Cities. CRC Press. 2018.

[16] Muthu B, Sivaparthipan CB, Kumar RL. Trust-based energy-efficient protocol over MANET using PTORA and RRFO. Wireless Pers Commun. 2024.

[17] Punitha P, Sivaparthipan CB, BalaAnand Muthu, Lakshmana Kumar R. A policy-configured resource management scheme for AHNS using link reliability K-means clustering algorithm and Weibull distribution-based blue monkey optimization. Int J Commun Syst. 2024;37(12):e5850.

[18] Punitha P, Lakshmana Kumar R, Revathi S, Premalatha R, Aiswarya RS. Secured framework with a hash function-enabled keyword search in cloud storage services. Int J Coop Inf Syst. 2024;33(3):2450001.

[19] Lakshmana Kumar R, Jayanthi S, BalaAnand Muthu, Sivaparthipan CB. An automatic anomaly application detection system in mobile devices using FL-HTR-DBN and SKLD-SED K-means algorithms. J Intell Fuzzy Syst. 2024;46(2):3245-3258.