Exploring Machine Learning Algorithms for Robust Cyber Threat Detection and Classification: A Comprehensive Evaluation

Ravikumar Ch
Assistant Professor, Department of AI&DS,
Chaitanya Bharathi Institute of Technology,
Hyderabad,-India.
chrk5814@gmail.com

Burri Naresh
Assistant Professor, Department of CSE-CS, *CVR College of Engineering-*
Hyderabad-India
naresh.burri@gmail.com

P Laxmi Prasanna
Assistant Professor, Department of CSE,
TKR College of Engineering & Technology,
Hyderabad- India.
prasanna5a2@gmail.com

Nenavath Chander
Assistant Professor, Department of CSE (CS),
Malla Reddy University,
Hyderabad, India
chander.nenavath@gmail.com

Ediga Amarnath Goud
Assistant Professor, Department of IT,
Sreenidhi Institute of Science and Technology,
Hyderabad, India
amar4goud@gmail.com

P Raghavendar Prasad
Associate Professor, Department of IT,
Mallareddy Engineering College,
Hyderabad-India.
prasad.mtech17@gmail.com

2024 International Conference on Intelligent Systems for Cybersecurity (ISCS) | 979-8-3503-7523-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISCS61804.2024.10581226

Abstract-*In response to the escalating frequency and complexity of cyber threats, the imperative need to enhance cybersecurity measures is evident. This study explores the potential of machine learning (ML) algorithms in advancing threat detection and classification by automating the identification of security incidents. The abstract presents a thorough assessment of various ML algorithms, including decision trees, support vector machines, and neural networks, for their efficacy in detecting and categorizing cyber threats. The evaluation encompasses a diverse dataset featuring different cyber-attack scenarios and incorporates multiple features such as network traffic patterns, system logs, and user behavior. Performance metrics, such as training accuracy and testing accuracy, are employed to assess the effectiveness of each algorithm. Furthermore, the study investigates the impact of feature selection techniques and model optimization strategies on algorithm performance. The results underscore the capability of ML algorithms to accurately identify and categorize cyber threats, providing valuable insights into their strengths and limitations. This research contributes to the field of cybersecurity by facilitating the development of practical and robust ML-based solutions, ultimately reinforcing cyber defence mechanisms against evolving threats.*

Keywords: cyber security, machine learning algorithms, threat detection, classification, training accuracy, testing accuracy

1. Introduction

Cyber security is a crucial aspect of computer science that deals with the investigation of various cyber threats and their corresponding countermeasures. With the widespread reliance on the internet for accessing information across different domains, such as business, education, and entertainment, network attacks have become increasingly prevalent [1]. To mitigate these attacks, intrusion detection systems (IDS) and firewalls have been recommended as preventive measures. While firewalls filter, IDS monitor the network for incoming and outgoing packets based on established rules. IDS are generally considered more effective and secure compared to firewalls [2].

The current challenge faced by computer networks lies in the diversity of cyber-attacks. These attacks vary in their nature and impact, ranging from adware that is relatively harmless to phishing attempts that can lead to data theft or destruction. More destructive attacks include ransomware, which encrypts computer systems and demands a ransom, and denial-of-service attacks that target operating systems [3]. In response to these challenges, researchers and engineers are focusing on the development of intelligent systems for automated computer network intrusion detection.

The research aimed to demonstrate the effectiveness of the investigated approaches in creating an intelligent system capable of identifying multiple anomalies within a network. To achieve this, the study explored machine learning and deep learning classification algorithms, particularly focusing on their promising outcomes in unsupervised modes of operation [4]. Machine learning techniques, known for their ability to identify newly discovered breaches swiftly, are commonly employed in the development of network intrusion detection systems [5]. In dealing with large datasets, precise algorithms for clustering, classification, and prediction are required, and supervised machine learning methods such as K-Nearest Neighbor (KNN) and Naive Bayes are frequently utilized. Decision trees, valued for their precision, versatility, and simplicity, play a significant role in detecting anomalous and abusive patterns. Additionally, neural networks have seen widespread deployment for detecting anomalies and abuse patterns [6]. The success of artificial intelligence models relies on both accuracy and interpretability, making the use of machine learning and deep learning approaches essential [7].

Despite existing literature on intrusion detection in cyber security, further advancements are still required. The present study contributes to addressing these issues by introducing the following approaches:

a) A comprehensive examination of the literature on the use of numerical and image-based datasets to machine learning and deep learning models for intrusion detection.

b) Dataset preprocessing and balancing.

c) Machine learning models that are stacked and use several feature extraction methods.

d) Run an experiment to confirm the models offered.

2. Literature Review

This section describes several studies that use fundamental machine learning methods to analyze IoT traffic in order to defend IoT devices against cyberattacks.
Network profiling and machine learning were the main topics of Rose et al.'s [7] study on IoT security. To detect unauthorized network transactions and attempts to tamper with IoT devices, they developed a dataset and a model. With a 98.35% accuracy rate and a 98.35% false-positive alert rate on the Cyber-Trust test, their suggested anomaly-based intrusion detection system produced outstanding results.

A broad machine-learning technique to recognize IoT devices was developed by Ali et al. [8] in a different work. During the training phase, they used random forest and naive Bayes classifiers to extract 85 features from packet capture (.pcap) files and obtained 99% accuracy in identifying IoT devices. Seven various supervised learning methods for IoT intrusion detection was evaluated side by side by El-Sayed et al. [9]. With 94% accuracy on MobileNetv2 features, the SVM technique performed better than the competition and showed rapid and consistent training outcomes.

According to Le K-H et al. [10], IMIDS is an intelligent intrusion detection system designed for Internet of Things (IoT) devices. It employs a compact convolutional neural network to classify various cyber threats, achieving an average F-measure of 97.22%. The system's detection capabilities were further improved through additional training using input from an attack data generator.

In a study conducted by Islam et al. [11], shallow and deep learning-based intrusion detection systems (IDSs) were examined for IoT threat detection. The deep learning-based IDSs included decision trees, random forests, and support vector machines. The performance of each participant was evaluated using five datasets, revealing that machine learning IDSs outperformed shallow machine learning techniques in accurately identifying IoT risks, achieving an accuracy of 98.79%.

Overall, these studies show the value and promise of using machine learning approaches to strengthen IoT security and efficiently identify cyber risks in IoT systems.

2.1 Network Attacks and their types

Network attacks can be defined as attempts to gain unauthorized access to a corporate network with the intention of stealing information or causing harm. These attacks can be categorized as either passive or active [13]. Passive attacks involve intercepting the network and gathering sensitive data without altering the system. Examples of passive attacks include traffic analysis and the unauthorized publication of message content. On the other hand, active attacks involve unauthorized access, where attackers can modify, delete, encrypt, or decrypt data during the attack. Replay attacks, denial of service (DoS), message manipulation, repudiation, and masquerade are examples of frequent active attacks.

a) Intrusion Detection Systems (IDS) are crucial for identifying and classifying different types of attacks, whether passive or active. The following are some specific types of attacks that IDS can consider:

b) Denial of Service (DoS): In this type of attack, untrusted users flood the network with meaningless traffic, aiming to exhaust its resources and prevent legitimate users from accessing it. Examples of DoS attacks include Land, Back, and Mail Blood Smurf attacks.

c) Probe Attack: This attack involves using software or a program to monitor and collect information about network activities. Examples of probe attacks include Satan, Ipsweep, Mscan, Saint, and Nmap.

d) Remote to Local (R2L): In R2L attacks, a hacker uses specific devices to transmit packets while being restricted from accessing the device's authorized account. The attacker exploits vulnerabilities to gain unauthorized access. Examples of R2L attacks include Named, Phf, Send mail, and Guest.

e) User to Root (U2R): In U2R attacks, the attacker has already gained access to a user's account and is attempting to exploit their privileges. Examples of U2R attacks include Perl, Ps, Eject, and Ffbconfig.

Being aware of these different types of attacks and employing effective intrusion detection systems is vital for safeguarding corporate networks and protecting sensitive information from unauthorized access and potential harm.

3. Methods and Materials

3.1. Datasets

The Kyoto dataset, the UNSW-NB15 dataset, the KDD cup dataset, and the nsl-kdd dataset are the four datasets that have been used in this experiment. These datasets were selected because they are useful and well-organized, especially when talking about machine learning techniques. The fact that these datasets are openly available is another benefit of using them. These datasets are also well-known for being user-friendly and high-quality, which makes them appropriate for data analysis.

Among these datasets, the KDD Cup dataset contains nearly 4.9 million single-connection vectors, each having 41 attributes. These vectors can be classified as either normal or attacking based on their behavior. The KDD Cup dataset includes data on four different types of attacks.

The Service Denial attack is one of the attack types featured in the KDD Cup dataset. Where the target device's memory is overwhelmed, causing it to become unresponsive when a request is received. To protect against this attack, the recommended action is to turn off the device.

Another attack type in the dataset is the user-to-root attack, where a hacker with specific access to a device tries to take control of the router by exploiting security vulnerabilities. Various methods, such as phishing attacks, sniffers (packet controlling), or social engineering, can be employed to carry out this type of attack. It occurs when a hacker, who does not physically possess the target computer, transfers packets from the computer to the network system and exploits security vulnerabilities to gain access to the target machine. Such attacks aim to obtain information from computer networks by bypassing the system's security measures.

Based on its specifications and row count, the KDD dataset has been selected as the dataset for this experiment. These features provide valuable insights into network traffic patterns, communication protocols, and potential network security threats. Researchers and analysts can leverage this data set to develop machine learning models and algorithms for network monitoring, intrusion detection, and overall network performance optimization.

Table 1: Description of the Data Set –Network Traffic Data

S. No	Feature Name	Data Type	Description
1	Source IP	String	The network or device's source IP address is used to send network traffic.
2	Destination IP	String	The network traffic's destination device's or network's IP address.
3	Source Port	Integer	The port that the source device or network users to send network traffic.
4	Destination Port	Integer	The port that the target device or network uses to receive network traffic.
5	Protocol	String	The protocol used for the network communication (TCP, UDP, ICMP, etc.)
6	Packet Length	Integer	The length (in bytes) of the network packet.
7	Timestamp	Date Time	The timestamp indicates the date and time of the network traffic event.
8	Network Protocol	String	The high-level network protocol used for the communication (HTTP, FTP, DNS, etc.)
9	Flow Duration	Integer	The duration (in ms) of the network flow.
10	Bytes Transferred	Integer	The total number of bytes transferred in the network flow.

3.2. Machine Learning Algorithms:

Several machine learning techniques, including Logistic Regression, Support Vector Machine (SVM), and Gaussian Naive-Bayes, were used in this investigation. Both linear and non-linear relationships can be handled using these strategies. The Multilayer Perception Algorithm, Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN) Algorithm were also used as Artificial Neural Network (ANN) techniques. Other algorithms used in the experiment included Gradient Boosting, Decision Tree, Random Forest, Stochastic Gradient Descent, K- Nearest Neighbor, and ANN [19][20].

Logistic Regression, being one of these algorithms, was applied as a classification technique. Since there were more than two possible outcomes, a Multinomial Logistic Regression technique was employed. The experiment was conducted using Python programming. Multinomial Logistic Regression is a variation of Binary Logistic Regression specifically designed for scenarios with multiple outcomes.

3.3 Feature Selection

A crucial step in data preparation for machine learning is feature selection, where we identify the most relevant attributes and discard the less important ones. The significance of a feature is determined based on how well it predicts the target variable [13].

In this study, the chi-square feature selection strategy is employed, which is particularly effective for multiclass classification [14]. By applying the chi-square test [15], the best feature for the dataset is identified, which indicates the feature with the strongest relationship to the output class. The Chi-square test formula is expressed as follows:

$$X2=\text{€}(Oij-Eij)2/Eij$$

In this formula:

- Oij represents the observed frequency for a particular feature and class.
- Eij represents the expected frequency for that feature and class, which is calculated based on the assumption that the feature and class are independent.

3.4 Proposed Model

The proposed model aims to improve the efficiency of classification through the utilization of machine learning techniques [16]. The underlying principle of the algorithms employed is as follows: the data is initially divided into groups based on certain criteria, and with each iteration of the algorithm, additional rules are incorporated into the existing set of rules. This iterative process reduces misclassification. By combining all the weak classifiers, a robust classifier capable of identifying various types of attacks is formed. One significant advantage of the AdaBoost method is its ability to evaluate the net classification error at each stage of learning [17][18]. This evaluation provides valuable insights into the overall performance and effectiveness of the classifier. It allows for the continual refinement and enhancement of the model's classification capabilities.

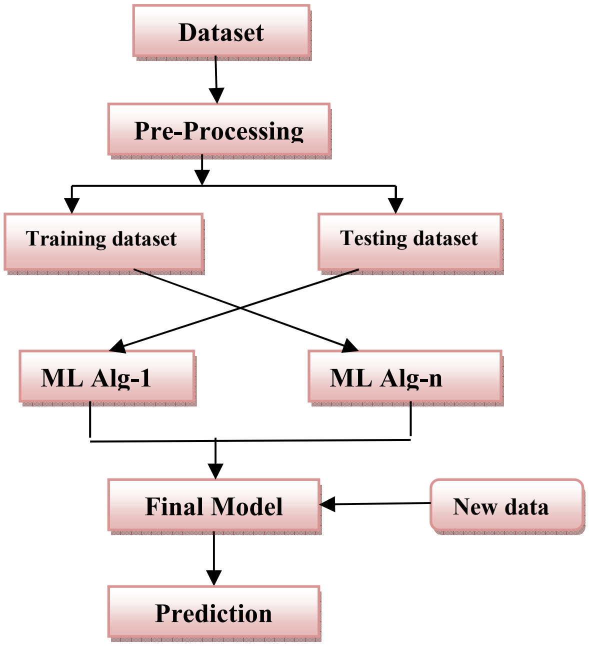

Figure -1: Proposed Model

3.5 Performance Metrics

Using the performance parameters described below [19], different IDS based on machine learning are compared and evaluated for their efficacy.

True Positive (TP) - Here, an assault has been acknowledged and confirmed to have occurred. Situations of this nature are seen positively.

Attacks that are mistakenly reported as having occurred are known as false positive (FP) attacks. A false positive is thus nothing more than a deceptive warning.

The data known as True Negative (TN) refers to those that can be rightfully categorized as normal and actually are normal. Such circumstances are seen to be quite awful. Details about attacks that have been mistakenly labeled as normal are known as false negative (FN) data. The most vulnerable stage is this one because no one is aware that an attack has already occurred.

The ratio of all observed values to the total of the TP and TN observations is used to measure accuracy. Accuracy is frequently predicated on the overall number of valid classifications. The accuracy formula is explained in the equation more fully.

$$\textbf{Accuracy} = TP + TN/TP + TN + FP + FN$$

4. Experiment & Results

The accompanying Tables 2 below show the outcomes of training and testing the suggested algorithms using the dataset. Both the attained testing times (s) and testing accuracies (%) are shown in the tables, together with the training times (s) and training accuracy percentages. The outcomes for the kddcup dataset are shown in Table 2.

These studies were conducted to demonstrate the effectiveness and accuracy of several kinds of intrusion detection algorithms as well as the processing time required to identify intrusions over the whole dataset. By demonstrating accuracy, we can evaluate them and help them learn on their own so that they may go on to perform with more precision in the future.

Table 2: Trail and Examination algorithm's results

Machine Learning Algorithms	Trail Period (s)	Examination period (s)	Trail Accuracy	Examination Accuracy
Gaussian Naïve-Bayes	0.9	1.0	58%	55%
Logistic Regression	1.3	1.0	96%	98%
SVM	1.8	1.6	90%	89%
Stochastic Gradient Descent	2.5	2.2	85	91%
Decision Tree Algorithm	1.2	0.9	98%	99%
Random Forest Algorithm	1.8	1.6	95%	96%
Gradient Boosting Algorithm	128	152	95%	95%
K-Nearest Neighbor	2.8	2.5	93%	88%
ANN	338	349	92%	91%
CNN	1243	1315	93%	95%
RNN	2145	2254	92%	93%

5. Conclusion and Future scope

As a result of their widespread use in businesses, machine learning and deep learning approaches have proven beneficial in mitigating cyber hazards. These methods enable businesses to automatically recognize, stop, recover from, and adapt to a variety of dangers without the need for explicit programming. A variety of algorithms were tested in the experiment, and while needing less development time, Logistic Regression and Decision Tree classifiers achieved extraordinary accuracy levels of over 95% in differentiating malware across test datasets. The accuracy of the Gaussian Naive-Bayes classifier ranged from 51% to 88%. Notably, the accuracy of the Random Forest Classification method fared better than all other algorithms. These findings show the usefulness of machine learning in successfully mitigating cyber attacks. Although decision trees and logistic regression are useful classifiers, the Random Forest Classification method was shown to be the most accurate of all the techniques examined.

Overall, these findings support continuing initiatives to strengthen cyber security and safeguard computer networks from new dangers. Machine learning algorithms will undoubtedly be employed more frequently in a number of industries in the upcoming years, including cyber security. The goal of this study was to identify malware using four distinct datasets and thirteen different classification techniques. Unexpectedly, 12 algorithms performed extremely well. It was discovered that simple to-construct and very powerful algorithms included the Gaussian Naive-Bayes classifier, Logistic Regression, and Decision Tree classifier.

6. References

[1] Sara Najari and Iman Lotfi, "Malware Detection Using Data Mining Techniques", International Journal of Intelligent Information Systems, Vol. 3, No. 6-1, December 2014, p. 33-37, DOI: 10.11648/j.ijiis.s.2014030601.16.

[2] Y. Qin and T. Xia, "Sensitivity analysis of ring oscillator-based hardware Trojan detection", 2017 IEEE 17th International Conference on Communication Technology (ICCT), 27-30 October 2017, Chengdu, China, pp. 1979-1983, ISSN: 2576-7828. DOI: 10.1109/ICCT.2017.8359975.

[3] Douglas Jacobson and Joseph Idziorek, Computer Security Literacy: Staying Safe in a Digital World, 1st ed. Florida, USA: Chapman and Hall/CRC, 27 November 2012, ISBN-13: 978-1439856185.-

[4] Dipanker Dasgupta, Zahid Akhtar and Sajib Sen, "Machine learning in cybersecurity: a comprehensive survey", The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology, Vol. 19, No. 1, 19 September 2020, pp. 57-16, DOI: 10.1177/1548512920951275.

[5] Rajashree A. Katole, Swati S. Sherekar and Vilas M. Thakare, "Detection of SQL injection attacks by removing the parameter values of SQL query", 2018 2nd International Conference on Inventive Systems and Control (ICISC), 19-20 January 2018, Coimbatore, India, pp. 736-741, DOI: 10.1109/ICISC.2018.8398896.

[6] Hafiz M. Farooq and Naif M. Otaibi, "Optimal Machine Learning Algorithms for Cyber Threat Detection", 2018 UKSim-AMSS 20th International Conference on Computer Modelling and Simulation (UKSim), 27-29 March 2018, Cambridge, UK, pp. 32-37, DOI: 10.1109/UKSim.2018.00018.

[7] Rose, J.R.; Swann, M.; Bendiab, G.; Shields, S.; Kolokotronis, N. Intrusion Detection using Network Traffic Profiling and Machine Learning for IoT. In Proceedings of the 2021 IEEE Conference on Network Softwarization: Accelerating Network Softwarization in the Cognitive Age, NetSoft 2021, Tokyo, Japan, 28 June–2 July 2021; pp. 409–415. [CrossRef]

[8]. Ali, Z.; Hussain, F.; Ghazanfar, S.; Husnain, M.; Zahid, S.; Shah, G.A. A Generic Machine Learning Approach for IoT Device Identification. In Proceedings of the 2021 International Conference on Cyber Warfare and Security (ICCWS), Islamabad, Pakistan,23–25 November 2021; pp. 118–123. [CrossRef]

[9]. El-Sayed, R.; El-Ghamry, A.; Gaber, T.; Hassanien, A.E. Zero-Day Malware Classification Using Deep Features with Support Vector Machines. In Proceedings of the 2021 Tenth International Conference on Intelligent Computing and Information Systems (ICICIS), Cairo, Egypt, 5–7 December 2021; pp. 311–317. [CrossRef].

[10]. Le, K.-H.; Nguyen, M.-H.; Tran, T.-D.; Tran, N.-D. IMIDS: An Intelligent Intrusion Detection System against Cyber Threats in IoT. Electronics 2022, 11, 524. [CrossRef].

[11]. Islam, N.; Farhin, F.; Sultana, I.; Kaiser, M.S.; Rahman, M.S.; Mahmud, M.; Cho, G.H. Towards Machine Learning Based Intrusion [8Detection in IoT Networks. Comput. Mater. Contin. 2021, 69, 1801–1821. [CrossRef].

[12] Subhash Waskle, Lokesh Parashar, and Upendra Singh. Intrusion detection system using pca with random forest approach. In 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), pages 803–808. IEEE, 2020.

[13] Shadi Aljawarneh, Monther Aldwairi, and Muneer Bani Yassein. Anomaly-based intrusion detection system through feature selection analysis and building a hybrid efficient model. Journal of Computational Science, 25:152–160, 2018.

[14] I Sumaiya Thaseen, Ch Aswani Kumar, and Amir Ahmad. Integrated intrusion detection model using chi-square feature selection and an ensemble of classifiers. Arabian Journal for Science and Engineering, 44(4):3357– 3368, 2019.

[15] Amar Meryem and Bouabid EL Ouahidi. Hybrid intrusion detection system using machine learning. Network Security, 2020(5):8–19, 2020.

[16] Nevrus Kaja, Adnan Shaout, and Di Ma. An intelligent intrusion detection system. Applied Intelligence, 49(9):3235–3247, 2019.

[17] Manjula C Belavagi and Balachandra Muniyal. Multiclass machine learning algorithms for intrusion detection performance study. In International Symposium on Security in Computing and Communication, pages 170–178. Springer, 2017.

[18] B Selvakumar and Karuppiah Muneeswaran. Firefly algorithm-based feature selection for network intrusion detection. Computers & Security, 81:148–155, 2019.

[19] Ansam Khraisat, Iqbal Gondal, Peter Vamplew, and Joarder Kamruzzaman. Survey of intrusion detection systems: techniques, datasets and challenges. Cybersecurity, 2(1):20, 2019.

[20] Ikram Sumaiya Thaseen and Cherukuri Aswani Kumar. Intrusion detection model using a fusion of chi-square feature selection and multiclass SVM. Journal of King Saud University-Computer and Information Sciences, 29(4):462–472, 2017.