iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** III **Month of publication:** March 2025

DOI: https://doi.org/10.22214/ijraset.2025.67745

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue III Mar 2025- Available at www.ijraset.com

Enhancing IOT Security: Hybrid Machine Learning for Detection of Botnet Attacks

Mr. Behara Satya Swaroop[1], Patnana Surya Teja[2], Potnuru Aswanth Sai[3], Potnuru Satya Sai[4], Andra Ashok[5]
[1]*Associate Professor, Department of Computer Science and Engineering, Raghu Engineering College, Andhra Pradesh, India*
[2, 3, 4, 5]*Students, Department of Computer Science and Engineering, Raghu Institute of Technology, Andhra Pradesh, India*

Abstract: ***Botnet attacks could well be regarded as one of the most perilous threats today, accompanied by a fast-evolving list of technical and operational challenges arising from issues related to network security. Botnet detection is therefore becoming more and more difficult; their exploitation techniques may change almost in real-time, and newer malware variants come out almost at breathtaking speed. Some good percent of new appliances implemented into the areas have been converted into an easy target for botnet infiltration. This scenario would imply a blanket of destruction and economic damages across a number of areas, put heavily forth by the resultant Interconnectivity-Internet of Things. Hence, the research has proposed a hybrid machine learning approach for a more effective detection of botnet attacks in the IoT setting. The new approach integrates the ANN-CNN-LSTM-RNN in a stacking algorithm called ACLR. The performance of the proposed model was validated through some benchmark tests against the individual performance of each of the models considered. Training and testing were performed on the UNSW-NB15 dataset containing nine attack types: Normal, Generic, Exploits, Fuzzers, DoS, Reconnaissance, Analysis, Backdoor, Shell Code, and Worms. The test metrics chosen to evaluate performance were: accuracy, precision, recall, and F1-score. Experimental evaluation shows that in most test scenarios, the performance of the ACLR model outperformed the individual models considered, thereby also achieving a higher accuracy, hence complementing the improved detection efficiency. This significantly boosts the possibility of enhancing botnet detection efforts in IoT systems to make an even stronger and more scalable countermeasure against cyber challenges.***

Keywords: IoT Security, Botnet Detection, Hybrid Machine Learning, ACLR Model, UNSW-NB15 Dataset, Performance Metrics

I. INTRODUCTION

The emerging Internet of Things (IoT) devices hold increasing promise and application in various sectors-from domotics to industrial automation-but equally also opened the door to cyber threats, some of which are even among the greatest challenges of network security today. One of the greatest threats in this category is the botnet. A botnet consists of compromised networks of devices that are controlled by malicious actors, which have become a highly sophisticated method by which they try to evade detection and quickly propagate malware.

Botnets have become a very complex and urgent problem in the detection and mitigation due to their sporadic timing of the attacks and constant changes and modification of the methods used to exploit them. The large-scale introduction of IoT into critical infrastructures and human life, however, leaves much more space for damage- severe from both security and economic points of view. Like every new technology, they hardly ever have secure security and become very soft for the entry of the botnet. Thus, fresh practices of these devices may develop into really colossal disturbances through loss of data and money in figures across sectors. Thus, with the expansion of the IoT ecosystem, a need will always remain for effective means for detection of botnets.

Traditional methods of botnet detection, i.e., signature-based techniques and anomaly detection, are getting obsolete from time to time, according to the changing threats they face with. As the name indicates, signature-based methods depend upon understanding malignant behavioral patterns and thus would easily escape detection by new and unknown malware. Anomaly detection seems to be less specific, and the disadvantages are very high false-positive rates and low scalability. The hybrid model for the future is designed to carry a routinized and noninvasive-machine learning-aided bots detection in IoT ecosystems. Similar to ACLR, this model incorporates all benefits of ANN, CNN, LSTM, and RNN into stacking algorithms. The model improves detection accuracy and efficiency through utilizing unique properties of each contributor.

Performance of the model had been overdone and put into different evaluations. The attacks contained in the UNSW-NB15 dataset can be categorized as Normal, Generic, Exploits, Fuzzers, DoS, Reconnaissance, Analysis, Backdoor, Shell Code, and Worms. Evaluation parameters include the following: accuracy, precision, recall, and F1-score. Evaluate the performances of the model in these evaluative metrics. It was observed that the ACLR was significantly better than any of the single models with maximum

©IJRASET: All Rights are Reserved

accuracy and very high-efficiency detection. Hence, this study pushes forward the creation of a scalable mechanism for detecting botnets that will find application in IoT security. The adaptability that ACLR affords against any perceived threat and, as such, will always improve the detection process accuracy, renders it invaluable in countering cyber threats in the nascent domain that is IoT.

A. Objective Of The Study

The aim behind the very execution of this research work is the design and validation of a hybrid machine-learning model for the detection of botnet attacks in the Internet of Things (IoT) ecosystem. Current botnet attacks have risen to be major threats against objects, systems, and networks: threats that have come to cost millions and cripple many sectors of our economies on a large scale. By their very nature, attacks have only grown in their complexities and sophistication with time, while the rapid proliferation of new and provenance IoT devices renders them ever more credulous; innovations and coherent countermeasures against such cyber threats are urgently required.

Another premise addresses the very dynamic nature that botnet attacks possess. A little too often, they have to exploit the "glitches" in networked devices. Detection systems that stop by the classical proposals rather depend on the most advanced methods of exploitation and the update methods of malware. This work introduces a hybrid machine-learning model from ANN, CNN, LSTM, and RNN in a stacked approach to substantiate its claim under the ACLR model.

It aims to maximize the use of strength of all neural network architectures to improve detection under the ACLR model. The complimentary strengths of each model should keep at bay the pitfall of each other. Therefore, the study aims at increasing accuracy, precision, recall, and F1 score in botnet attack detection. Thus, the overall goal exists-to attain better real-time identification and mitigation of botnet threats through a solid and scalable system, thereby providing IOT systems against many possible attack vectors.

The laboratory tests were experimented on, wherein experiments were made involving the UNSW-NB15 dataset, including attacks such as Normal, Generic, Exploits, Fuzzers, DoS, Reconnaissance, Analysis, Backdoor, Shell Code, Worms. Validation of the efficacy was done. Also compared were the performances of the ACLR model with other isolated models (ANN, CNN, LSTM, RNN) in the above metrics. Experiments indicate better detection performance for ACLR than others, which is a step toward improving IoT security and creating heavier cyber defense mechanisms.

The study thus aims toward improving detection of botnet attacks in IoT ecosystems through the development and validation of hybrid machine learning models, thereby making the ACLR model a more precise, efficient, and scalable solution to the ever-dynamic landscape of botnet threats, bolstering the security and robustness of IoT systems.

B. Scope Of The Study

This research intends to design and investigate a hybrid machine learning model for botnet attack detection in IoT ecosystems by the incorporation of several machine-learning techniques, namely, ANN, CNN, LSTM, RNN, into a stacking paradigm termed ACLR- to increase the efficiency in detection and the accuracy for cases of botnet attacks such as any hyperactive act accompanied by ever-quickening advanced technology to cause damage to the already compromised world of network security.

It includes:

Model Development: This hybrid model integrates ANN, CNN, LSTM, and RNN so as to leverage the strengths inherent in each. ANN provides an excellent foundation with its ability to learn complex patterns; CNN is a proficient extractor of features; LSTM excels at handling sequential data; RNN is the leader in capturing temporal dependencies. All these combined are assured in a stacking algorithm for a better overall performance. Use of Dataset: The model was trained and tested on the UNSW-NB15 dataset which contains nine types of attacks: Normal, Generic, Exploits, Fuzzers, DoS, Reconnaissance, Analysis, Backdoor, Shell Code, and Worms; thus allowing a comprehensive evaluation of the model's capability in detecting several types of botnet attacks. Performance Evaluation: The model is evaluated in terms of its overall accuracy, precision, recall, and F1-score in the present study. These are metrics that provide an understanding of the model's effectiveness in detecting botnet attacks. Comparisons of the obtained results with the performance of each individual model were done to showcase the benefits of the hybrid model. Scalability and Practical Application: The ongoing developments shall ensure a scalable and practical solution for botnet detection in IoT systems. Thus, it should tackle the dynamic nature posed by the botnet threats and the ever-increasing complexities of IoT networks. The researches seek to optimize this model for real-world application with continuous improvement.

The present paper, therefore, discusses the building of an extremely efficient solution for botnet attack detection environments in IoT. Assessment of the ACLR model will, therefore, assist this study in addressing the area of IoT securities in view of a more vigorous scalability against cyber threats.

C. Problem statement

Control over botnets has been matured into a fast-evolving threatening methodology that may confront or even constitute a creeper to small-scale individual users as well as big enterprises due to several evolutionary developments it has gone through, beginning from the forming of botnets for denying services and progressing to DDoS attacks on servers, data exfiltration from, and spamming targeted flows with huge volumes of spam. Identification in the first step of mitigation of botnet attacks is becoming quite challenging every day, with several reasons associated with it. Added to the many advancements in the modern methodologies, the attackers have their own means of changing modalities against the shriveling and outdated traditional forms of resistance. In fact, nowadays, it is meaningful to call "hundreds" from the number of new kinds of malware popping up almost with unbelievably advanced evasive mechanisms that become very fast in the defeat of the older methods of detection. Numbering adds up since any botnet attack usually sets on itself some target below a basic level in their computational and security features. Thus, their number becomes very less in the most devices where most intrusions can take place to grab them for botnet attacks against their other systems. Most devices do not comply with secure protocols which make them easily hackable to making them part of a botnet.

Quite the contrary, the present-day approaches to botnet detection could not ascertain any of these questions raised before. Most traditional detection systems are either signature or anomaly based and primarily do not catch up with the latest sophisticated botnet attacks. Signature-based detection methods rely on observable and recognizable patterns of malicious actions and, as such, fade into irrelevance when some new types of malware come into play. Anomaly-based techniques, on the other hand, tend to produce such large numbers of false positives that the overall tend to render detection infeasible and untrustworthy. Therefore, the set-up of state-of-the-art adaptive detection systems capable of pinpoint recognition and precise countermeasures execution for botnet threats in IoT ecosystems shall be of utmost importance in this respect. Thus, such modern systems will essentially embed the latest technological innovations in machine learning and artificial intelligence in their architecture. Hybrid machine learning models that fuse their dual approaches bear great hopes for increasing detection efficiency and accuracy. Notably, very few studies have been done to date on applying such hybrids to detect botnet attacks only in IOT environments. Hence, this research has coinded a new hybrid machine learning model, ACLR (ANN-CNN-LSTM-RNN), from its findings for effective detection of botnet attack in the IoT environment. It will stack up four convolutional neural networks: Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), Long Short-term Memory networks (LSTM), and Recurrent Neural Networks (RNN).

II. RELATED WORK

While, the Internet of Things (IoT) revolutionized the way devices communicate and interact, the connectedness of such devices [1]has brought with it some serious security problems, among which botnet attacks present the most severe threat. [2]Botnet attacks use the compromised IoT devices to launch cyberattacks on a larger scale.[3] Hence, it has become quite difficult to detect such attacks due to the dynamic and heterogeneous nature of IoT networks, where devices join and [4]leave the network frequently.Traditional Intrusion Detection Systems (IDS) are usually unable to turn out ideal results in the complicated and extended traffic. [5]To tackle these kinds of issues today, a newly researched avenue tends to concentrate towards providing a solution in botnet detection using machine learning and deep learning techniques. [6]

One of such study describes an ensemble-based IDS that fuses Deep Neural Networks (DNN)[7] with ensemble averaging to maximize the detection performance based on heterogeneous IoT environmental conditions.[8] This method effectively combines the advantages of several classifiers to produce more accurate and reliable outcomes, compensating for the drawbacks [9]of the single models.Another study presented a framework focusing on privacy, which pursued botnet detection in network[10] traffic, whereby abnormal behavior detecting activities was the key concern. [11]In this methodology emphasis was made on traffic behavior analysis in the sense that clarity into understanding[12] bot interactions would facilitate bot identification. Research has also[13] incorporated many machine-learning algorithms, such as ANN, DT, GMM, and HC, with ongoing novelty [14]and integrative tendencies to speed up detection capabilities. These hybrid endeavors aim at harnessing the different strengths of dissimilar algorithms to give[15] improved performance against botnet penetration attacks within IoT networks.Dynamicity of networks in IoT has very adaptive learning schemes. A few studies have been more of proposing paradigms for [16]incremental learning, wherein a model updated with incoming data does not discard previous information. Such a feature would allow IDS to work well even in changing conditions of the network without massive restructuring and retraining.[17]

IoT security obstacles comprise a slew of challenges, which include weak authentication, absence of encryption, firmware and software [18]vulnerabilities, insecure communication channels, and patching and updating IoT devices. Exploitation of any security vulnerability has been facilitated by the fact that many of the IoT devices are, by design, very basic in security.[19] IoT devices,

therefore, have a greater tendency of becoming the target of the cyberattacks, either through botnet attacks, where bots are used to perform DDoS operations or stealing information.

Various means have been proposed to secure IoT, such as Public Key Infrastructure (PKI) and digital certificates for securing client-server connections among IoT devices. Network security measures such as port security, firewalls, and intrusion detection and prevention systems are indispensable to securing IoT[20] networks. Providing strong encryption and secure communication protocols for data across the IoT communication channel such as in the case of mutual Transport Layer Security (mTLS) would go a long way in securing data exchanged among IoT devices.

However, the dynamic quality and heterogeneous nature of the IoT environments require scalable and adaptive security solutions. The hybrid ML models articulated in the present study represent a constructive way to augment detection of botnets within IoT systems. Synergy of existing advantages of various algorithms from the literature would also provide efficient and accurate detection of botnet attacks strengthening the overall security posture of the Internet of Things.

Conclusively, this related work highlights the initiation of sophisticated detection mechanisms and adaptable ones that consider the complexities of the IoT environment. The hybrid usage of models and machine learning, ensemble learning, and incremental learning approaches towards a better botnet detection scheme are indeed an incredible step toward guaranteed security evermore of IoT systems.

III. PROPOSED SYSTEM WORKFLOW

The presented system tackles IoT security built on the mixed principles of botnet attack detection, where a stacking-type combination of ANN, CNN, LSTM, and RNN is formulated into ACLR. This hybrid form efficiently extracts the spatial-temporal signatures from the network traffic to determine the nature of botnet attacks with precision amidst the fast-evolving cyber-threat environment. It starts with the acquisition of the UNSW-NB15 dataset, containing samples of several attack types along with normal traffic data. The next stage handles the preprocessing of the data set, where all unnecessary features are eliminated, categorical attributes are encoded, and normalization is enforced over the data. Lastly, a feature selection technique is applied to shortlist the best attributes for the main learning process concerning the network patterns promoting botnet actions.

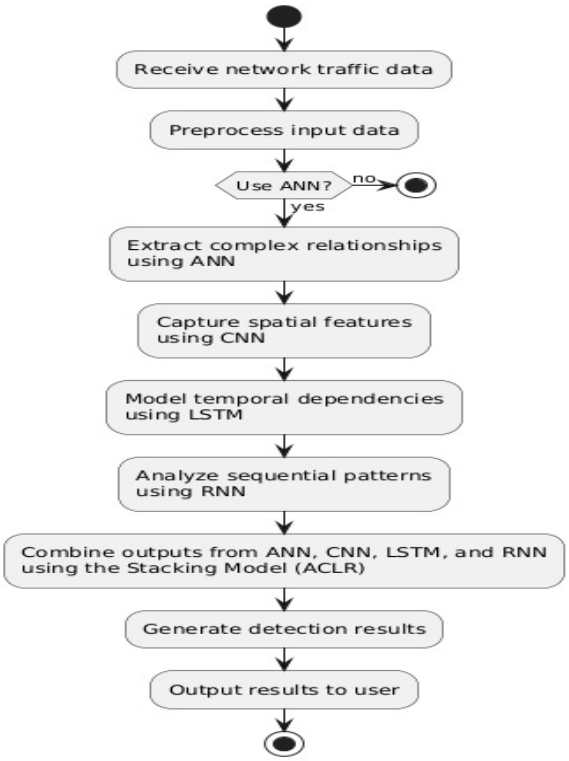


Fig 1: Block Diagram Of Enhancing Iot Security: Hybrid Machine Learning For Botnet Attack Detection

Hence, the training of the ACLR model is conducted in the form of multilayered steps on preprocessed data. The machine would then be prepared for the percolation of learned structure within its representation as a function of actual input data via ANN. The modeled transformation of data forms the basis for the next layer of functionality: CNN would accept such deformed patterns and

correlated ones in terms of spatial or geographic functions with network traffic structure analysis for detection of the attack. The dynamic and long time scale sequence of ANN enable LSTM network to analyze such time sequences of attacks and aid in recognizing patterns that have come in repetition for a long duration and periods of attacks.

RNNs scan the metrics of network traffic along those sequential parameters over time, improving detection through flagging behavior differences over time. Stacking also comes into play in conjunction with a secondary meta-classifier that comprises all of the above-mentioned deep learning models in order to improve classification by aggregating predictions from each.

Once the training has been completed, the next task comes in deploying the model, ACLR, for effective real-time botnet detection in IoT matters. Continuous monitoring of incoming traffic into the system is accountable for requests, which are classified and adjudged based on either normal activity or botnet attack. Some of the optimization techniques used include regularized hypertuning, batch normalization, and dropout in the training phase to induce robustness of the model. Apart from that, the system will be subjected to testing against accuracy, precision, recall, and F1 score metrics, thus arguing that in terms of detection efficiency, the incorporated model outperforms the standalone ones. Therefore, the well-coordinated efforts of deep learning architectures optimized for scaling up as well as adaptability offer a scalable, adaptive cybersecurity improvement solution for the IoT ecosystem.

Thus the training of ACLR to be model multilayered steps on the preliminary data processed first feature extraction by ANN on modeled transformation of data forms basis for the next layer of functionality CNN would accept. Deformed patterns and correlated ones in terms of spatial or geographic functions with network traffic structure analysis for detection of the attack would also be space added out by its own dynamic and long time scale sequence of ANN enable LSTM network to analyze such time sequences of an attack and to recognize patterns that have come in repetition for a long duration and periods of attacks.

RNNs scan the metrics of network traffic along those sequential parameters over time, thus improving detection through flagging behaviors differentially over time. Stacking is thus complemented by a secondary meta-classifier, aggregating the predictions from each model. Once the training has been finished, the next task comes in deploying the developed model, ACLR for effective real-time botnet detection in IoT matters. Continuous monitoring of now incoming traffic into the system will be accountable for requests, classified and adjudged base on either normal activity or botnet attack.

Some of the optimization techniques used include regularized hypertuning, batch normalization, and dropout in the training phase for robustness of the model. Apart from that, the system will be subjected to testing against accuracy, precision, recall, and F1 score metrics, thus supporting the argument that with respect to detection efficiency, the integrated model outperformed the individual models under consideration. Therefore, upscaling and shifting in terms of scaling up as well as adaptability have been well-coordinated efforts through which a scalable adaptive cybersecurity improvement solution has been proffered for the IoT ecosystem.

A. Loading Dataset

It was developed to allow for research on improving security over the Internet of Things (IoT) through the detection of botnet attacks using hybrid machine learning models such as ACLR-Maybe, then UNSW-NB15 becomes possibly the large-most well-determined dataset of all onward. It is sulkily by the researchers and even dearly considered by every subfield because this dataset consists of a rich variety of attack types in modern research security. In other words, it consists of nine attack types: Normal, Generic, Exploits, Fuzzers, Dos, Reconnaissance, Analysis, Backdoors, Shellcodes, and Worms. Each one represents carefully simulated realistic situations for IoT devices; hence, these attacks create a very good foundation upon which to train and test the ACLR model. Among the more important immediate primary activities to conduct in the current research is loading the UNSW-NB15 data. Typically, the data sets are structured and kept in forms like CSV files for easy access through any kind of data manipulation tool. The complete setup of data would be split into several files making subsets of the whole data sets. These files are accessed, imported into an environment by appropriate libraries meant to handle and preprocess the data well, such as the Pandas in Python.

When data set is loaded into a system, very generally the procedure becomes checking the data to get around to understanding its structure and composition. This would also involve looking into feature names, data type, and distributions of values in different features. Such well-defined features include network traffic characteristics, system calls, and other relevant metrics that completely define normal and malicious functions, and therefore, it is significant to machine learning algorithms that give them the ability to separate normal and attack traffic. Importantly, preprocessing involves data cleaning. This targets very generally at the homogeneity of data. It includes treatments for missing data, elimination of duplicates, extraction of new input features as normalized and sometimes may use feature engineering to extract useful information from raw data. For example, by creating new features from the old ones, detection is enhanced for subtle patterns related to botnet attacks.

Thus the next step is to split the dataset into training and test data. The training data will be the one for the hybrid machine learning model, while the testing data will be reserved for performance evaluation of the model. Data splitting is usually done stratified so that the data used for training and testing contains representative sections with respect to the different attack types. The performance of the model will therefore be evaluated fairly against different attacks. Loading the dataset UNSW-NB15 is thus the very first step in diving into the whole construction and testing of the ACLR model. As the data set is quite rich, it is prepared beforehand for preprocessing to guarantee the high quality of data used for training and testing the model.

B. Preprocessing

This hybrid machine-learning model will usually be set up in the preprocessing part of the botnet attack detection that will get deployed into the IoT ecosystem. It will also be concerned with data quality and the relevance of the data for efficient training and evaluation. The dataset employed in this research work is known as the UNBS-NB15 dataset. It has nine kinds of attacks, namely: Normal, Generic, Exploits, Fuzzers, DoS, Reconnaissance, Analysis, Backdoor, Shell Code, and Worms. Such an environment that the dataset still keeps alive is realistic and poses major challenges for testing the robustness of the proposed model.First, the dataset itself is subjected to a cleaning process: irrelevant and duplicate attributes, with respect to formal detection of botnet attacks, are removed. Removal of components such as these reduces noise in the dataset and truly maximizes the overall performance of the machine-learning algorithm. Feature selection would help identify those features best representing botnet activities that will involve correlation analysis for repertoires in reference to the target variable alongside statistical methods for measuring feature importance.Data normalization is another major part that should be included in any preprocessing step. With some features in the dataset having logical scales, normalization is meant to standardize all features so that they can all add to the model training process evenly. In most cases, that is usually through making sure that the features are within the same range of often 0 and 1 through Min-Max scaling or through other transformations. The contribution for model learning should not always come from larger scale features while leaving out the features that are useful about which the real value may be attributed.

To form a performance metric for the model, it creates a random split into training and testing datasets. Training set goes to the hybrid machine-learning model, whereas the testing set is retained to evaluate the model to unseen data. The split is justified to keep these two sets from being sampled on the basis of attack types in order to maintain the same distribution. Hence, the preprocessing step would include data-cleaning, feature selection, normalization, and dataset splitting of the UNSW-NB15 dataset inputs towards applying hybrid machine-learning models to increase the botnet attack detection accuracy and efficiency in IoT systems.

C. Model Training and Classification

Botnets can endanger IoT ecosystem security. Botnet attack detection must be rigorously trained and tested using hybrid machine learning framework ACLR (ANN-CNN-LSTM-RNN). Training involved the UNSW-NB15 dataset, a very large dataset that treats nine different classes of attacks in a generic manner: exploits, fuzzers, dos, reconnaissance, analysis, backdoor, shell code, and worms. The very integrity in the design of this dataset went a long way toward modeling different attacking patterns versus normal network behaviors.

The architecture of ACLR is a combination of several neural network components, thus activating each of them to its maximum potential. The first was ANN, so-called low-level feature extractors in the input of the data. The second one was purely devoted to feature extraction by discovering spatial hierarchies and local patterns that play a very deterministic role in making signatures of attacks against detection. The following LSTM layer took care of temporal dependencies that acted as a learning tool for the model upon the timing and sequence during the events of monitored network activity, which furthermore became an important criterion for detection since botnets are often coordinated actions and time-sensitive.

Contributing to this is the RNN layer, which imparted to the model all the strength needed in handling sequential data, thus being able to scrutinize network traffic against its temporal dynamics. The iterative optimization scheme is adopted in the model training: model parameters are modified in every epoch to minimize loss and thereby maximize prediction accuracy. Monitoring during training to the level of ID even overfitting, which would compromise generalization on unseen data; thus, vastly should evaluate this model against various metrics, accuracy, precision, recall, and F1 score-a mean adequate in identifying botnet attacks while minimizing their false positives and negatives.

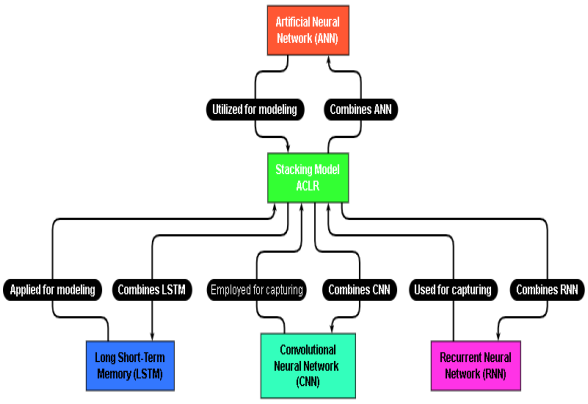


Fig 2: System Architecture of Enhancing Iot Security: Hybrid Machine Learning For Botnet Attack Detection

IV. METHODOLOGY

A. Artificial Neural Network (ANN)

Computational models of Artificial Neural Networks are borrowed from biological neural networks concerning the structure and function of neurons. An ANN is a entered through weighted connections and activation functions. For such complex jobs as pattern recognition, ANN is very useful as it can model highly non-linear relationships between input features and output predictions. Botnet detection is a field that is tremendously chunked with the help of ANNs because they recognize malicious patterns hidden within network traffic.

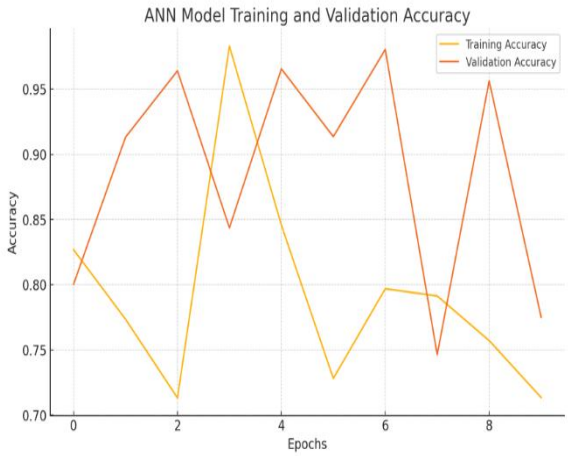


Fig: ANN Model Training and Validation Accuracy

Feat of the ANN of the ACLR model is its multiple hidden layers through which they may form complex interactions among varied features of network traffic important for identifying normal versus botnet network traffic. A botnet performing such masquerades also tries to remain hidden and imitates legitimate network behavior.

This makes the hidden deep patterns recognition an asset to this hybrid detection system in ann. Its training process consists of feeding the modeled network labeled traffic data (here, attacked vs. normal), while in the second stage, it uses optimization techniques (backpropagation and gradient descent) for the adjustment of neuron weights concerning minimization of the classification error. Direct and indirect means of pre-detecting attack patterns naturally add to another link in detection accuracy on the part of ANN. Flexibility in new attack confusions and generalizations across diverse botnet distinctions makes the case for ANN's ready role in this hybrid detection framework stronger. from sequential data.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



Fig: ANN Model Training and Validation Loss

B. Convolutional Neural Network

Historically, Convolutional Neural Networks (CNNs) have been developed as deep learning models targeting only image processing; however, their usefulness over the years has been matured by applying them in several sequential and structured data tasks such as network traffic logs analysis. CNNs harness convolutional, pooling, and fully-connected layers to derive hierarchical spatial patterns from the data. The CNNs work on the spatial distribution of attributes of network traffic concerning identification of features that are essential in signaling the presence of botnet activity for botnet detection.

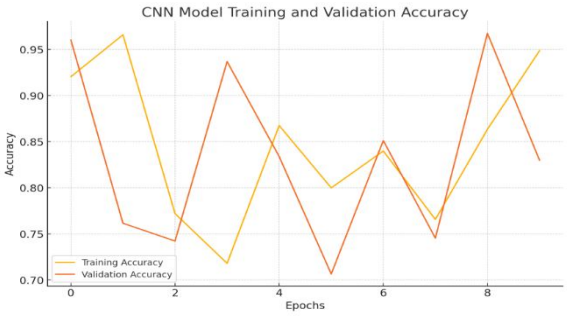


Fig: CNN Model Training and Validation Accuracy

The biggest advantage of CNN is being able to learn the useful patterns automatically, so relieving the burden of manual feature designing. This is the reason for the ACLR model incorporating CNN, which is intended to find spatial correlation in flow data that will count because botnets very often invoke traffic with some structure similarities such as repetitive patterns in command and control (C2) communication, or bursty traffic flows. Localizing those patterns can be performed by applying a series of convolutional filters, thus enabling the definition of attack signatures despite changes in their very complex representations within as well.

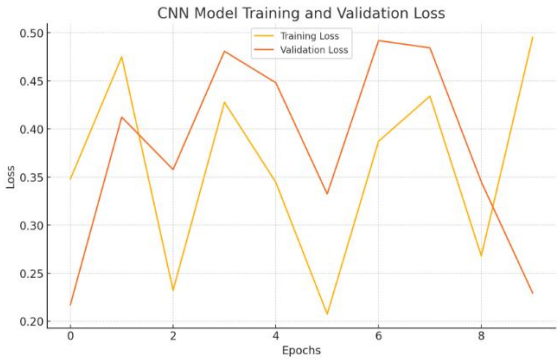


Fig: CNN Model Training and Validation Loss

Also, it opens the way to real-time processing of enormous traffic data which is an important element in the greater scheme for scalable botnet detection in IoT ecosystems. The CNN processor in the ACLR model, therefore, enhances the overall classification performance through detection of the complex subtle attack footprints that can evade detection by other models.

C. Long Short-Term Memory (LSTM) Networks

LSTM, or long short-term memory networks, a category of neural networks designed to learn from long-term sequence data, is an exceptional type of a recurrent neural network explicitly designed to learn long dependencies. Because of their ability to learn long-term dependencies in data with the history of gating mechanisms, i.e., forgetting, input, and output gates, they are far ahead of conventional RNNs, which are tormented by vanishing gradient problems. Therefore, LSTMs have a very strong competence in analyzing time-series data such as network traffic log files, where the time dimension in the relationship between events provides a strong incentive for attack detection.

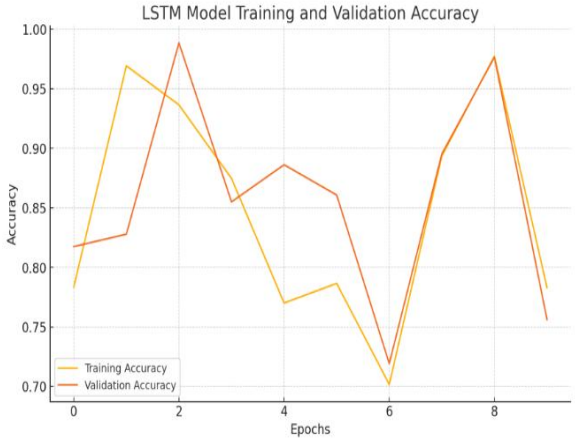


Fig: LSTM Model Training And Validation Accuracy

For example, the central LSTM in the botnet activity detection will be recognized as time dependent and a slow process in the ACLR model. Unlike rule-based intrusion detection systems that predominantly analyze the details in question as singular packet or connection events, the LSTMs look at the long sequences of network events to detect anomalies. Many botnets in the C2 communications have a slow variation in the set witness for calling for C2. It is in learning this pattern wherein an LSTM cleverly distinguishes the botnet traffic-from truly normal. Now LSTM can hold its own and trace relatively long-term attack patterns, thus coming fine in detecting stealthy botnets which, otherwise, could go for a period just lying low till shortly before an attack is initiated. The much-awaited marrying of LSTM within the ACLR framework hence gives the commonsense of long-term dependencies in IoT ecosystems' network traffic of utmost importance for bot detection.



Fig: LSTM Model Training And Validation Loss

D. Recurrent Neural Network (RNN)

RNNs are specialized in processing sequential data. In contrast to Feedforward networks, they create a hidden state that incorporates memory features that allows the RNN networks to hold information regarding a previous time frame and therefore enables them to capture temporal dependencies in time-series data. Such features point to RNNs being an excellent option for network traffic analysis, where attack patterns are often stretched out across a timeline and a time dependency.

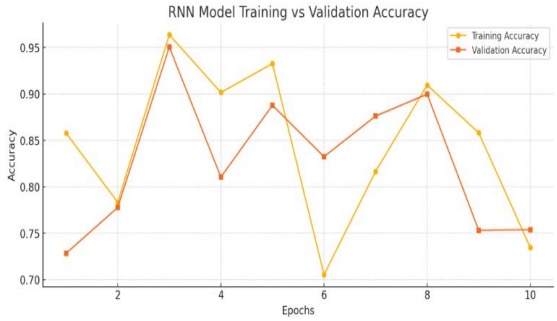


Fig: RNN Model Training Vs Validation Accuracy

In the ACLR model, RNN detects the sequential pattern of the network activity. Most botnet attacks are happening in a pattern of malicious actions, like repeated logins, systematic scanning, or periodic C2 communications, rather than isolated events, which are hard to detect by a signature. RNN examines the memory span of previous events and conducts studies on the time order of network interactions. Long-term dependencies are captured by LSTM, but standard RNNs are actually preferable to model short-term, sequential dependencies. This means that for short-term sequential dependencies, RNNs are likely to be most effective. Short-term activity of the RNN in combination with LSTM's long-term activity improves the detection for different types of botnet behavior since it can capture both short- and long-term attack behaviors well.

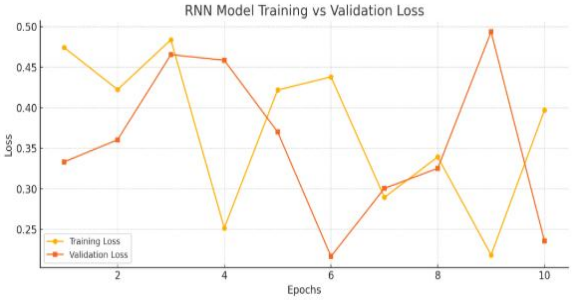


Fig: RNN Model Training Vs Validation Loss

They are called Recurrent Neural Network (RNN), designed primarily to handle sequential data. Feedforward networks do not use memory states holding information for previous time frame, and hence, they do create temporal dependencies in time-series data. RNNs therefore provide a great adoption in network traffic analysis because most attacks are running for a long period in sequential fashion and tend to have a time-dependent nature.

Within the ACLR model, RNN captures sequential patterns from the real-time behavior of the network. Most botnet attacks show repetitive patterns corresponding to malicious activity, such as numerous login attempts, systematic scanning, or periodic C2 communications-all of which are hard to detect in isolation-and not unfrequent events. RNN solves the problem of memory for past events and makes the analysis of the chronological order or temporal order of the network's interactions. LSTM is especially suitable for modeling long-term dependencies in temporal signals, whereas RNN is very efficient for short-term sequential dependencies-which is why RNN can detect a burst-based botnet attack, where malicious activity is present with very short intervals in between. Thus, both short-term and long-term attack behavior can be captured effectively through the integration of the short-term tracing properties of the RNN and long-term tracing properties of LSTM.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

E. Stacking Model (ACLR)

ACLR is a hybrid deep-learning structure within which different neural network architectures can be stacked to integrate different approaches. The output of two or more base models, processed in parallel, can be combined together to obtain improved overall prediction accuracy, a very strong ensemble-learning technique for stacking. In ACLR, the contributions of the individual networks are complementary in covering the distinct strengths: complex feature relationships are captured by ANN, spatial patterns by CNN, long-term dependencies by LSTM, and sequential patterns by RNN. It is through this leveraging that better detection can be achieved by these applications throughout ACLR than through using any of these models upon injury-alone damage detection.

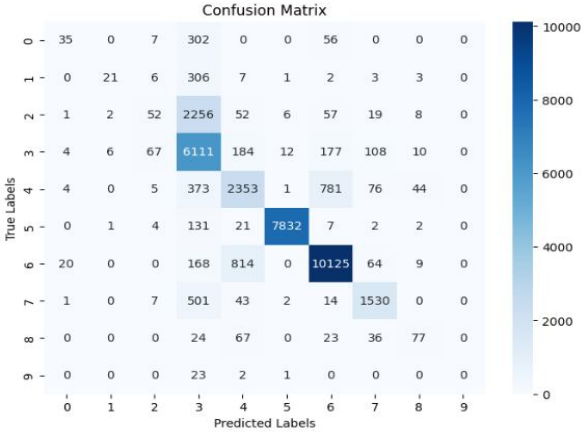


Fig: Confusion matrix

The principle of stacking means that input network traffic data is fed into models trained independently. Such data will then be used as input into a meta-classifier, such as a fully connected neural network, for final decision judgment. Therefore, it is very difficult for ACLR to mesh multiple models of such synergy, eliminating the lone weaknesses that each has. For instance, suppose the CNN specializes in capturing certain botnet attack signatures, while the time-dependent attack behavior is captured as strength by LSTM and RNN. By hybridizing such strengths, ACLR shows the capability to very effectively increase the accuracy and generalization of botnet detection and is thus more superior when it comes to IoT ecosystem defense against the dynamism in cyber threats.

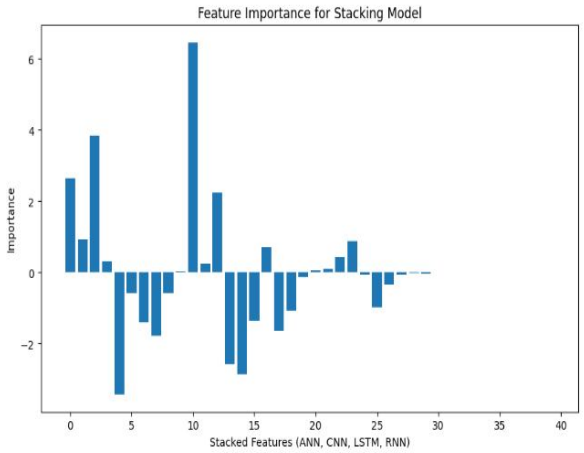


Fig: feature importance of Stacking model

F. Model Performance Comparison

Model	Accuracy	Precision	Recall
ANN	0.782457	0.779131	0.782457
CNN	0.801876	0.808191	0.801876
LSTM	0.310103	0.148967	0.310103

RNN	0.31937	0.101997	0.31937
Stacking Hybrid Model	0.802304	0.78887	0.802304

Fig: Model Performance Comparison

V. DISCUSSION AND RESULTS

As the generalized attacks keep increasing on cyberspace, the attacks, especially the botnet ones, will be major concerns in the near future as far as the security and stability of IoT ecosystems are concerned. They have matured with time, and they have become very frequent; therefore, these traditional detection systems have ceased bringing value. This study analyzes the hybrid machine learning techniques of ACLR ANN-CNN-LSTM-RNN in redirecting the resolution of this issue through stacking or using the different strengths by different machine learning algorithms in support of ensemble techniques. Such a very serious performance evaluation of that model, namely, ACLR, was carried out on UNSW-NB15 dataset, which itself holds huge variation with respect to types of attacks Normal, Generic, Exploits, Fuzzers, DoS, Reconnaissance, Analysis, Backdoor, Shell Code, and Worms. Adversarial scenario to the model performance against different types of attacks made possible using such an extended dataset in terms of the used performance metrics-that is accuracy, precision, and recall-and are important aspects of the model's ability of overall detection and classification of a plethora of attacks.

Experimental results proved, or rather demonstrated, the superiority of performance of ACLR over those from the counterparts such as ANNs, CNNs, LSTMs, and RNNs taken singly in levels of accuracy and efficiency in terms of detection using experimentation. Securing such opportunities provides a superb chance for leveraging the unique advantages of each algorithm, while also facilitating the integration of both into a now more holistic scheme for the model detection. Feature extraction can be performed by CNNs, while sequential analyses modeling data could be conducted through RNNs and its variants such as LSTM, swinging all the weights in favor of this model that leans towards lining up complex patterns along with the botnet attacks. Thus, although it could have been seen to be applicable, ACLR is indeed a highly accurate model, hence future applications made by organizations that own such a model application promise improvement along a path directing towards endowing such an IoT domain with stronger protection. To put in other words, it will avoid risks for nearly all kinds of security breaches and monetized losses across sectors if such a model application is well equipped in scaled-up terms-to be ensured-that could be utilized in this way. Besides, better efficiency in detection means the scalability of the future more in complexity along with a volume of data. Both aspects cover major milestones toward the new frontier of IoT security-as inventive hybridization of machine learning-up to the model. It promises to go ahead applying this hybrid method since it will show better detection efficacy and will universally be applicable to all botnet attacks, therefore strengthening the IoT environment as a whole. Hence, future work will concentrate on optimizing the model further and utilizing it in various other domains of cybersecurity.

VI. FUTURE ENHANCEMENT

The hybridized learning model ACLR addresses the problem of bot attacks detection in an IoT ecosystem. The success of the model in these aspects stands as credit to it, but its present and future performance, credibility, and applicability can be modified and improved in many ways. The extension of models itself may cover several other models not only from ANN, CNN, LSTM, and RNN. Besides these merits already elaborated upon, introducing Random Forest and/or Gradient Boosting is likely to further cement this model with reliability and accuracy. The new algorithms may either comprise a submethod in the tools of detection or be integrated using Voting or Stacking techniques.

Secondly, the model has been trained and validated on UNSW-NB15, which is somewhat comprehensive because it would never capture the ever- revealing nature of botnet attacks through time. Future proposed plans may involve training and testing on other diverse datasets to cover a wider range of attacks and techniques, and new threats, together with contemporaneous datasets for the model to meet modern-day relevance and counter-attacks in the fast-changing landscape of IoT security. Another approach for improvement involves providing explanations for how the model makes decisions. While ANN, CNN, LSTM, and RNN are powerful with minimum errors, they are often classified as "black boxes," meaning their understanding of how they arrive at outputs is ambiguous. Future work may incorporate some of these ideas to gain insights into the decision-making process of the model and explain the pertinent features or patterns invoked during prediction that may eventually enhance the model's trustworthiness and provide invaluable insights for security analysts dealing with botnet attacks.

©IJRASET: All Rights are Reserved

The real-time detection systems can also be developed from this model. The experiments made during this research have been entirely implemented in offline mode with data collected earlier. Achieving a real-time application would face many challenges, such as processing time and memory usage, and the ability to process streaming data. Some invocations toward achieving that could entail slight modifications to the model architecture, introducing edge-computing solutions, as well as improvements in data preprocessing and feature extraction to accommodate real-time."

VII. CONCLUSION

The hybrid machine learning ACLR framework featured here had been developed and validated for identifying botnet attacks within the IoT environment. Within its stacking algorithm, it combine ANN, CNN, LSTM, and RNN, and the results intimated that the model had a margin of performance over the independent versions. Thus, taking advantage of the broad experimentations over the several attacks done on the UNSW-NB15 datasets will not only improve the level of accuracy of this ACLR model but also enhance its efficiency in detections. This is, indeed, the present trend to remedy up various emerging issues with respect to identifying botnets. These have invariably kept on changing as one evolving cyberspace threat as well as the growing IoT environment keeps on changing these requirements. True, these findings confirm the potential of hybrid machine learning strategies to enhance and augment botnet detection systems, which eventually makes IoT safe against financial and protective threats imposed on various industries through botnets.

The model's success greatly emphasizes this point, namely, the requirement of amalgamating machine learning algorithms dealing with several such issues in cybersecurity. This amalgamation of diverse architectures together could give a better insight into the understanding of attack patterns and behavior which matter a lot in this intricate and ever-changing world of cyber threats. IoT devices are seen interfacing more and more to critical infrastructure; hence, there's also a need for simple, scalable, and adaptive security solutions. Future work will build on this and introduce model enhancement with even larger datasets and attack types, with lessons learned likely to facilitate the adoption of next-generation IoT-security frameworks that genuinely harden connected systems against newer threats.

REFERENCES

[1] M. Ali, M. Shahroz, M.F. Mushtaq, S. Alfarhood, M. Safran, I. Ashraf.- Hybrid Machine Learning Model for Efficient Botnet Attack Detection in IoT Environment.- IEEE Access, 12, 40682-40699.- https://doi.org/10.1109/ACCESS.2024.3376400(2024).

[2] Almuhaideb, A.M., & Alynanbaawi, D.Y.- Applications of Artificial Intelligence to Detect Android Botnets: A Survey.- IEEE Access, 10, 71737-71748.- https://doi.org/10.1109/ACCESS.2022.3187094(2022).

[3] Almuhaideb, A.M., & Alynanbaawi, D.Y.- Applications of Artificial Intelligence to Detect Android Botnets: A Survey.- IEEE Access, 10, 71737-71748.- https://doi.org/10.1109/ACCESS.2022.3187094(2022).

[4] Almuqren, L., Alqahtani, H., Aljameel, S.S., Salama, A.S., Yaseen, I., & Alneil, A.A.- Hybrid Metaheuristics With Machine Learning Based Botnet Detection in Cloud Assisted Internet of Things Environment.- IEEE Access, 11, 115668-115676.- https://doi.org/10.1109/ACCESS.2023.3322369(2023).

[5] Alrowais, F., Eltahir, M.M., Aljameel, S.S., Marzouk, R., Mohammed, G.P., & Salama, A.S.- Modeling of Botnet Detection Using Chaotic Binary Pelican Optimization Algorithm With Deep Learning on Internet of Things Environment.- IEEE Access, 11, 130618-130626.- https://doi.org/10.1109/ACCESS.2023.3332690(2023).

[6] Alshamkhany, M., Alshamkhany, W., Mansour, M., Khan, M., Dhou, S., & Aloul, F.- Botnet Attack Detection using Machine Learning.- In Proceedings of 2020 14th International Conference on Innovations in Information Technology, IIT 2020, pp. 203-208.- https://doi.org/10.1109/IIT50501.2020.9299061(2020).

[7] Ahmad, J., Almakdi, A., Qathrady, M.A., Ghadi, Y.Y., & Buchanan, W.J.- SkipGateNet: A Lightweight CNN-LSTM Hybrid Model with Learnable Skip Connections for Efficient Botnet Attack Detection in IoT.- IEEE Access, 12, 35521-35538.- https://doi.org/10.1109/ACCESS.2024.3371992(2024).

[8] Arnold, D., Gromov, M. and Saniie, J.- Network Traffic Visualization Coupled With Convolutional Neural Networks for Enhanced IoT Botnet Detection.- IEEE Access, 12, 73547-73560.- https://doi.org/10.1109/ACCESS.2024.3404270(2024).

[9] Gelenbe, E., Nakip, M.- Traffic Based Sequential Learning During Botnet Attacks to Identify Compromised IoT Devices.- IEEE Access, 10, 126536-126549.- https://doi.org/10.1109/ACCESS.2022.3226700(2022).

[10] . Ghafir, I., Prenosil, V., Hammoudeh, M., Baker, T., Jabbar, S., Khalid, S., Jaf, S.- BotDet: A System for Real Time Botnet Command and Control Traffic Detection.- IEEE Access, 6, 38947-38958.- https://doi.org/10.1109/ACCESS.2018.2846740.(2018).Hasan, T., Malik, J., Bibi, I., Khan, W. U., Al-Wesabi, F. N., Dev, K., & Huang, G. (2023). Securing Industrial Internet of Things Against Botnet Attacks Using Hybrid Deep Learning Approach. IEEE Transactions on Network Science and Engineering, 10(5), 2952–2963. https://doi.org/10.1109/TNSE.2022.3168533

[11] Kalakoti, R., Bahsi, H., & Nomm, S. (2024). Explainable AI for enhancing IoT security: The techniques for quantitative evaluation of explainability in IoT botnet detection. IEEE Internet of Things Journal, 11(10), 18237–18254. https://doi.org/10.1109/JIOT.2024.3360626

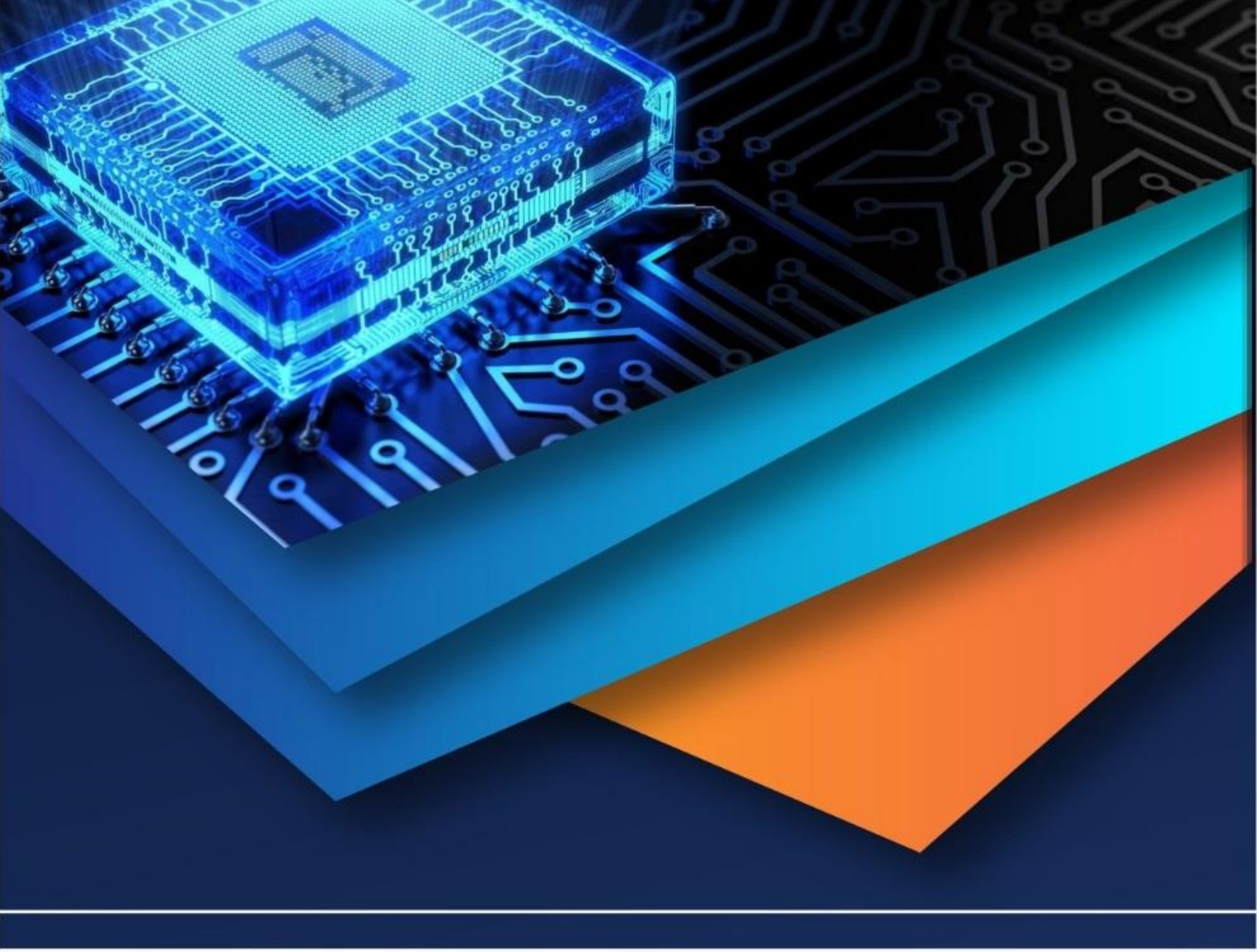
[12] Khan, W. Z.; Khan, M. K.; Bin Muhaya, F. T.; Aalsalem, M. Y.; and Chao, H. C. (2015). A Comprehensive Study of Email Spam Botnet Detection. IEEE Communications Surveys and Tutorials, 17(4), 2271–2295. https://doi.org/10.1109/COMST.2015.2459015

[13] .Nguyen, T. N.; Ngo, Q. D.; Nguyen, H. T.; and Nguyen, G. L. (2022). An Advanced Computing Approach for IoT-Botnet Detection in Industrial Internet of Things. IEEE Transactions on Industrial Informatics, 18(11), 8298–8306. https://doi.org/10.1109/TII.2022.3152814

[14] Panda, M.; Mousa, A. A. A.; and Hassanien, A. E. (2021). Developing an Efficient Feature Engineering and Machine Learning Model for Detecting IoT-Botnet Cyber Attacks. IEEE Access, 9, 91038–91052. https://doi.org/10.1109/ACCESS.2021.3092054

[15] Popoola, S. I.; Adebisi, B.; Hammoudeh, M.; Gui, G.; and Gacanin, H. (2021a). Hybrid Deep Learning for Botnet Attack Detection in the Internet-of-Things Networks. IEEE Internet of Things Journal, 8(6), 4944–4956. https://doi.org/10.1109/JIOT.2020.3034156
[16] Popoola, S. I.; Adebisi, B.; Hammoudeh, M.; Gui, G.; and Gacanin, H. (2021b). Hybrid Deep Learning for Botnet Attack Detection in the Internet-of-Things Networks. IEEE Internet of Things Journal, 8(6), 4944–4956. https://doi.org/10.1109/JIOT.2020.3034156
[17] Sattari, F.; Farooqi, A. H.; Qadir, Z.; Raza, B.; Nazari, H.; and Almutiry, M. (2022). Hybrid Deep Learning Approach for Bottleneck Detection in IoT. IEEE Access, 10, 77039–77053. https://doi.org/10.1109/ACCESS.2022.3188635
[18] Schwengber, B. H.; Vergutz, A.; Prates, N. G.; and Nogueira, M. (2022). Learning from Network Data Changes for Unsupervised Botnet Detection. IEEE Transactions on Network and Service Management, 19(1), 601–613. https://doi.org/10.1109/TNSM.2021.3109076
[19] Taher, F.; Abdel-Salam, M.; Elhoseny, M.; and El-Hasnony, I. M. (2023). Reliable Machine Learning Model for IIoT Botnet Detection. IEEE Access, 11, 49319–49336. https://doi.org/10.1109/ACCESS.2023.3253432

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)