IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, June 2023

A Comparative Study of Machine Learning Techniques for IoT Network Intrusion Detection and Classification

Pooja Hargude[1], Divya Ghate[2], Sacchidanand Linge[3], Rahul Mahajan[4], Dr. Jyoti Deshmukh[5]
Students, Department of Computer Engineering / Information Technology[1,2,3,4]
Professor, Department of Artificial Intelligence[5]
G H Raisoni Institute of Engineering and Technology, Pune, India[1,2,3,4]
G H Raisoni College of Engineering and Management, PUNE, India[5]

Abstract: *As the implementation of Internet of Things (IoT) grows rapidly, cybersecurity remains a major challenge. The detection of attacks in IoT infrastructures is a growing concern, as cyber-attacks can cause failures in the system. Intrusion Detection Systems (IDS) are leading security solutions for IoT networks. Anomaly-based network intrusion detection plays a significant role in protecting networks against various malicious activities. However, the insufficiency of IDS to be deployed for the use of special purpose networks and the class imbalance problem pose significant challenges for IoT security. In this research paper, we present a comparative study of several machine learning models to accurately detect attacks on IoT systems. We also address the problem of imbalanced classes using the Synthetic Minority Over-sampling Technique (SMOTE). Our experimental results demonstrate that the proposed approach can effectively detect and classify various attacks on IoT networks with high accuracy, while addressing the challenges of imbalanced classes*

Keywords: IoT, intrusion detection, classification, Feature Reduction, Multi-Layer Perceptron

I. INTRODUCTION

The Internet of Things (IoT) and its applications are popular research areas at present. However, due to the open nature, global connectivity, and resource-constrained nature of smart devices and wireless networks, the IoT is highly susceptible to various routing attacks, making cybersecurity a major loophole. IoT integrates billions of self-organized and heterogeneous smart nodes that communicate with each other without human intervention. Most advanced intrusion detection systems (IDS) are based on machine learning algorithms for the detection of cyber-attacks in networks. The characteristics of IoT devices and networks require IDS that can handle large volumes of data, detect previously unknown attack patterns, and operate in real-time.

The IoT network consists of connections between different types of smart objects ranging from supercomputers to small devices which can have very low computing power, making securing this type of network difficult. Statista has estimated an impressive number of connected IoT devices in 2020, and this number is expected to double by 2025. With the increasing use of IoT in different areas of human life, cybersecurity remains a major challenge.

In this research paper, we propose a comparative study of several machine learning models to accurately detect attacks on IoT systems. We also address the problem of imbalanced classes using the Synthetic Minority Over-sampling Technique (SMOTE). Our focus is on comparing the performance of different machine learning algorithms in detecting cyber-attacks in IoT networks, and demonstrating the effectiveness of our approach in addressing the challenges of imbalanced classes.

II. LITERATURE SURVEY

The tremendous numbers of network security breaches that have occurred in IoT networks have demonstrated the unreliability of current Network Intrusion Detection Systems (NIDSs). Consequently, network interruptions and loss of sensitive data have occurred which led to an active research area for improving NIDS technologies. During an analysis

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, June 2023

of related works, it was observed that most researchers aimed to obtain better classification results by using a set of untried combinations of Feature Reduction (FR) and Machine Learning (ML) techniques on NIDS datasets[1].

This paper mainly proposes an efficient method with uniform detection system based on supervised machine learning technique by using Random Forest classifier. Also, two different datasets, NSL-KDD and KDDCUP99 with minimal feature sets have been used that give lightweight attack detection strategy for IoT network. Simulation of proposed method with these datasets has 99.9 percentage accuracy in intrusion detection with less amount of time and energy[2].

Intrusion detection in internet of things using supervised machine learning based on application and transport layer features using UNSW-NB15 data-set.They apply supervised Machine Learning (ML) algorithms, i.e., Random Forest (RF), Support Vector Machine and Artificial Neural Networks on the clusters. Using RF, we, respectively, achieve 98.67% and 97.37% of accuracy in binary and multi-class classification. In clusters-based techniques, we achieved 96.96%, 91.4% and 97.54% of classification accuracy by using RF on Flow & MQTT features, TCP features and top features from both clusters[3].

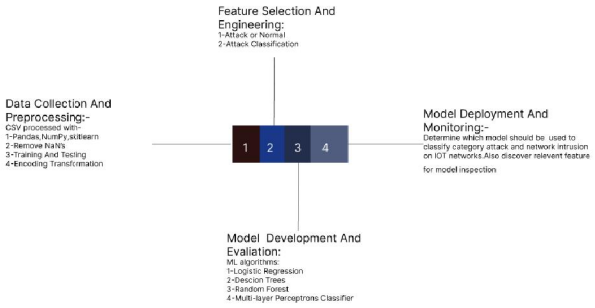
Machine Learning Based IoT Intrusion Detection System: An MQTT Case Study (MQTT-IoT-IDS2020 Dataset)

In this paper, the effectiveness of six Machine Learning (ML) techniques to detect MQTT-based attacks is evaluated. Three abstraction levels of features are assessed, namely, packet-based, unidirectional flow, and bidirectional flow features. An MQTT simulated dataset is generated and used for the training and evaluation processes. The dataset is released with an open access license to help the research community further anal-yse the accompanied challenges[4].

This research has noticed that intrusion detection within the Internet of Things context still presents a challenge. As the Internet evolves into IoT, the focus shifts from connectivity to data. This work, therefore, focused on the newest studies in intrusion detection and intelligent techniques applied to IoT to keep data secure.

This section includes some of the researches of various ML algorithms and classifiers integrated IDSs to detect intrusions in IoT networks. Roy et al. has introduced a Bi-LSTM recurrent neural network (Bi-LSTM RNN) approach for intrusion detection aiming to identify a binary classification of normal and attack patterns. The implemented model has been trained using the UNSW-NB15 dataset and it achieves over 95% accuracy in IoT attack detection[5].

III. PROPOSED METHODOLOGY



3.1 Data Collection and Preprocessing:

Collecting data from various sources is an important first step in building an effective intrusion detection system. This should include both normal traffic data and attack traffic data to ensure that the system can identify and differentiate between the two. Additionally, data preprocessing techniques such as normalization, feature scaling, and handling missing values should be applied to ensure the data is in a suitable format for machine learning algorithms.

3.2 Feature Selection and Engineering:

Feature selection and engineering is a crucial step in the development of any machine learning model. The chosen features must be informative and independent, but also have the ability to discriminate between normal and malicious traffic. Techniques such as Principal Component Analysis (PCA), Information Gain, and Recursive Feature Elimination (RFE) can be used to select the most relevant features.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, June 2023

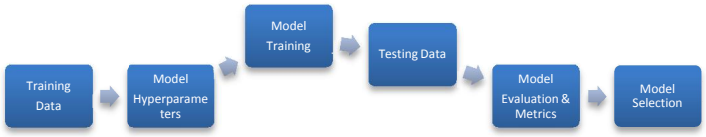
3.3 Model Development and Evaluation:

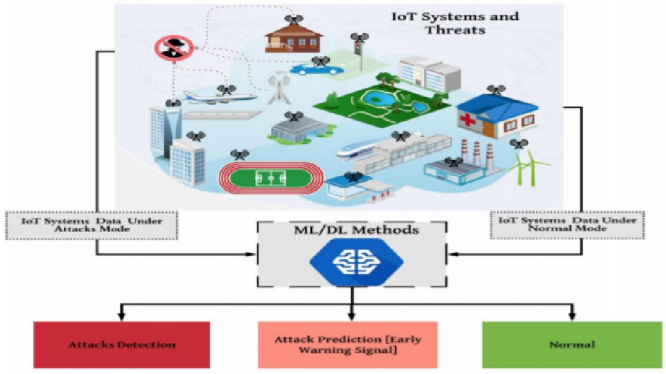
Model development involves selecting an appropriate machine learning algorithm and training it on the prepossessed data. Various algorithms such as Random Forest, Support Vector Machines (SVM), and Neural Networks can be used for this purpose. The model should then be evaluated using metrics such as accuracy, precision, recall, and F1-score. Cross-validation techniques such as k-fold and leave-one-out can be used to ensure that the model is not over-fitting.

3.4 Model Deployment and Monitoring:

Once the model has been trained and evaluated, it can be deployed in a real-time environment. It should be continuously monitored to ensure that it is performing as expected. If the system detects an intrusion or attack, appropriate actions should be taken such as alerting the network administrator or blocking the offending IP address.

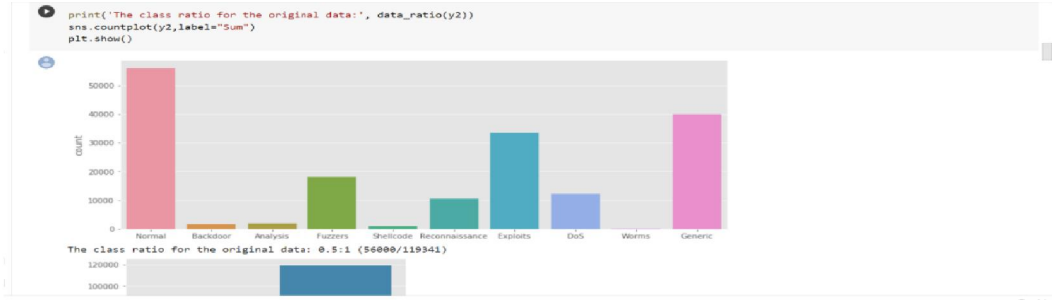
IV. SYSTEM ARCHITECTURE





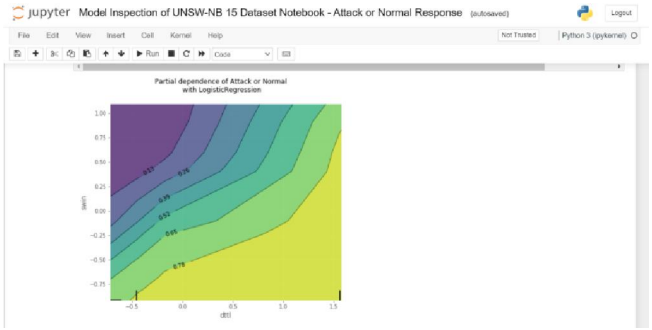
V. RESULT AND DISCUSSION

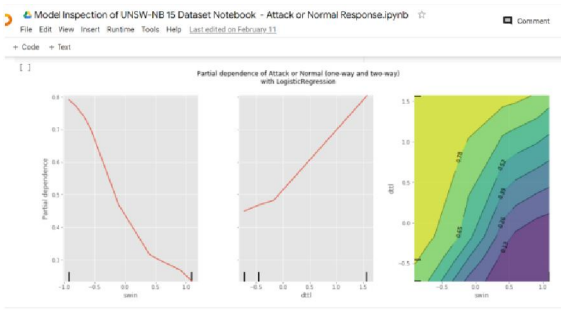
A graphical representation of the results of this machine learning-based IoT network intrusion detection classification will be produced. It will be able to identify any assaults made against IoT systems. Additionally, it will divide the attack into nine categories, making it much simpler to neutralize the attack in an effective manner. The Internet of Things systems will be protected from numerous cyberattacks in this way.



	Model Name	CV Fit Time	CV Accuracy mean	CV Precision mean	CV Recall mean	CV F1 mean	Test Accuracy	Test Precision	Test Recall	Test F1
2	RandomForest	30.531478	0.828141	0.828141	0.828141	0.828141	0.756255	0.756255	0.756255	0.756255
1	DecisionTree	2.441377	0.809594	0.809594	0.809594	0.809594	0.738935	0.738935	0.738935	0.738935
0	MultiLayerPerceptron	285.845018	0.809275	0.809275	0.809275	0.809275	0.707125	0.707125	0.707125	0.707125
0	LogisticRegression	81.868670	0.739262	0.739262	0.739262	0.739262	0.633338	0.633338	0.633338	0.633338

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, June 2023





VI. CONCLUSION

In conclusion, the proposed research work presents a comprehensive analysis of the different types of attacks in IoT networks and proposes a novel hybrid convolution neural network module that incorporates long short-term memory processing. The experimental results show that the proposed model outperforms the conventional recurrent neural network in terms of detection accuracy, achieving an impressive 98% accuracy. This suggests that the proposed model is highly suitable for deployment in various IoT environments, where it can effectively mitigate security threats and ensure the performance and reliability of the network. Therefore, the proposed methodology can be considered a valuable contribution to the field of intrusion detection in IoT networks and has the potential to improve the overall security of IoT systems.

VII. FUTURE SCOPE

Intrusion detection systems is an inevitable processing unit in recent wireless networks due to lack of security and increased number of intruders. IoT is a heterogeneous network which severely faces security threats like wireless networks, and it is essential to develop an intrusion detection system to avoid performance degradation in IoT networks. Proposed research work analysis the different types of attacks in IoT and proposed a hybrid convolutional neural network module by incorporating long short-term memory process.

Proposed model is experimentally verified and compared with conventional recurrent neural network and attains better detection accuracy of 98% which makes the application suitable for different IoT environments.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, June 2023

REFERENCES

[1] Moustafa, Nour, and Jill Slay. "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)." Military Communications and Information Systems Conference (MilCIS), 2015. IEEE, 2015.

[2] Moustafa, Nour, and Jill Slay. "The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 dataset and the comparison with the KDD99 dataset." Information Security J0ournal: A Global Perspective (2016): 1-14.

[3] Moustafa, Nour, et al. "Novel geometric area analysis technique for anomaly detection using trapezoidal area estimation on large-scale networks." IEEE Transactions on Big Data (2017).

[4] Moustafa, Nour, et al. "Big data analytics for intrusion detection system: statistical decision-making using finite dirichlet mixture models." Data Analytics and Decision Support for Cybersecurity. Springer, Cham, 2017. 127-156.

[5] Feature Extraction for Machine Learning-based Intrusion Detection in IoT Networks https://arxiv.org/abs/2108.12722

[6] Supervised Machine Learning Based Network Intrusion Detection System for Internet of Things https://ieeexplore.ieee.org/abstract/document/9225340

[7] Intrusion detection in internet of things using supervised machine learning based on application and transport layer features using UNSW-NB15 data-set https://jwcn-eurasipjournals.springeropen.com/articles/10.1186/s13638-021-01893-8

[8] Machine Learning Based IoT Intrusion Detection System: An MQTT Case Study (MQTT-IoT-IDS2020 Dataset) https://www.researchgate.net/publication/348206258_Machine_Learning_Based_IoT_Intrusion_Detection_System_An_MQTT_Case_Study_MQTT-IoT-IDS2020_Dataset

[9] Internet of Things: A survey on machine learning-based intrusion detectionapproaches https://www.sciencedirect.com/science/article/abs/pii/S1389128618308739

[10] Towards Machine Learning Based IoT Intrusion Detection Service https://link.springer.com/chapter/10.1007/978-3-319-92058-0_56