A Comparative Analysis of Machine Learning Techniques for Anomaly Detection in IoT Networks Using Secondary Data

Ayush Ransingh
Department of Computer Application
Haribhai V. Desai College, Pune, India

Abstract — The rapid expansion of Internet of Things (IoT) devices has introduced significant security challenges due to their limited computational capabilities and increased exposure to cyber threats. Detecting anomalies in IoT networks has become essential for maintaining system integrity and preventing unauthorized activities. This paper presents a comparative analysis of various machine learning techniques for anomaly detection using secondary data derived from existing research studies. Techniques such as Random Forest, Decision Tree, Naïve Bayes, Support Vector Machine (SVM), and Neural Networks are evaluated based on performance, computational complexity, and suitability for resource-constrained environments. The analysis highlights that lightweight models provide a practical balance between efficiency and detection accuracy, whereas deep learning techniques offer improved performance at the cost of higher resource consumption. The study also emphasizes the importance of selecting appropriate models based on system requirements and available computational resources.

Keywords: Internet Of Things (IoT) Security; Anomaly Detection; Machine Learning; Intrusion Detection Systems; Lightweight Models; Network Security;

I. INTRODUCTION

The Internet of Things (IoT) has transformed modern technological systems by enabling seamless communication between connected devices such as sensors, smart appliances, and industrial machines. This interconnected environment has improved automation, efficiency, and real-time monitoring across various domains. However, the rapid growth of IoT has also increased the risk of cyber threats.

Most IoT devices operate with limited memory, processing power, and security mechanisms. These constraints make them highly vulnerable to cyberattacks such as Distributed Denial of Service (DDoS), spoofing, and unauthorized access [1]. Traditional security approaches rely on predefined rules and signatures, which are not effective in detecting modern and evolving threats.

Machine learning techniques have emerged as a promising solution for anomaly detection in IoT networks. These techniques can learn patterns from network traffic and identify deviations that indicate malicious activity [2]. In addition, IoT systems generate large volumes of structured data, making them suitable for machine learning-based analysis.

The objective of this study is to analyze and compare different machine learning techniques for anomaly detection in IoT environments using secondary data from existing research studies.

II. LITERATURE REVIEW

Several research studies have explored the application of machine learning techniques in IoT anomaly detection. These studies emphasize the importance of using realistic datasets that reflect actual network behaviour and attack patterns.

Tree-based models such as Decision Trees and Random Forest are widely used due to their ability to handle complex datasets efficiently. These models provide high accuracy while maintaining relatively low computational complexity, making them suitable for IoT environments [3].

Naïve Bayes is another commonly used algorithm due to its simplicity and efficiency. It performs well in scenarios where quick predictions are required, although its assumptions may limit its effectiveness in complex datasets.

Support Vector Machines (SVM) have been shown to provide high classification accuracy in structured datasets. However, their computational requirements make them less suitable for real-time IoT applications [2].

Deep learning models such as Neural Networks offer advanced pattern recognition capabilities and can detect complex attack patterns. Despite their advantages, these models require significant computational resources and are typically deployed in centralized systems.

Recent studies also highlight the importance of data preprocessing and feature selection in improving model performance [4]. These steps help reduce noise and improve the efficiency of anomaly detection systems.

III. METHODOLOGY

This research adopts a secondary data analysis approach, where data from previously published studies is analyzed instead of implementing new machine learning models. Relevant research papers were selected based on their focus on IoT anomaly detection and machine learning techniques.

The selected studies were analyzed to extract key parameters such as algorithms used, datasets, and reported performance metrics. These parameters were then organized to enable a structured comparison of different techniques.
The evaluation criteria used in this study include:

- Accuracy
- Computational complexity
- Resource usage
- Suitability for IoT environments

The comparison is conducted based on reported results from literature, ensuring consistency and reliability in analysis.

IV. RESULTS AND DISCUSSION

Technique	Accuracy	Complexity	Resource Usage	Suitability
Random Forest	High	Medium	Moderate	Highly Suitable
Decision Tree	Medium-High	Low	Low	Suitable
Naïve Bayes	Medium	Low	Very Low	Suitable
SVM	High	High	High	Moderate
Neural Network	Very High	Very High	Very High	Less Suitable

Table I: Comparison Of Machine Learning Techniques

The results indicate that Random Forest provides an optimal balance between performance and efficiency. Decision Trees and Naïve Bayes are suitable for resource-constrained environments due to their low complexity and faster execution.

SVM offers high accuracy but requires greater computational power, making it less suitable for edge-level IoT devices. Neural Networks achieve the highest accuracy but demand significant processing resources.

The findings suggest that the selection of a machine learning model depends on the specific requirements of the IoT system. Lightweight models are more suitable for edge devices, while complex models can be deployed in centralized systems.

It is also observed that the performance of machine learning models depends on the quality of the dataset. Proper preprocessing and feature selection can significantly improve model efficiency and detection accuracy.

V. ADVANTAGES AND LIMITATIONS OF ML TECHNIQUES IN IOT

A. Advantages

Machine learning techniques enable detection of unknown and evolving cyber threats by analyzing behavioral patterns rather than relying on predefined rules. These models can adapt over time as new data becomes available, improving detection accuracy.

They also reduce the need for manual monitoring and enable automated anomaly detection, making them suitable for large-scale IoT systems.

B. Limitations

The performance of machine learning models depends on the availability of high-quality datasets. In IoT environments, collecting labelled data can be challenging.

Resource constraints in IoT devices limit the deployment of complex models. Deep learning techniques, although accurate, are not always feasible for real-time applications. Issues such as overfitting and false positives can also impact system reliability.

VI. FUTURE SCOPE

Future research can focus on developing hybrid models that combine the strengths of multiple machine learning techniques. The use of real-time datasets and continuous learning approaches can improve detection accuracy.

Integration of cloud computing and edge computing can help overcome resource limitations and enable the use of advanced models in IoT environments.

VII. CONCLUSION

This study presented a comparative analysis of machine learning techniques for IoT anomaly detection using secondary data. The findings indicate that lightweight models such as Random Forest and Decision Tree are most suitable for practical IoT environments.

While deep learning models offer higher accuracy, their computational requirements limit their applicability. Selecting an appropriate model requires balancing performance and resource constraints.

REFERENCES

[1] Moustafa, N., & Slay, J. (2016). Evaluation of Network Anomaly Detection Systems using UNSW-NB15 dataset.
[2] Doshi, R., Apthorpe, N., & Feamster, N. (2018). Machine Learning DDoS Detection for IoT Devices.
[3] Hindy, H., et al. (2020). Intrusion Detection Systems for IoT: A Survey.
[4] Shafiq, M., et al. (2021). IoT Network Traffic Analysis using Machine Learning.
[5] Ferrag, M. A., et al. (2020). Cyber Security for IoT: Attacks and Countermeasures.

