Impact of Ransomware on IoT Network Security: Analysis and Defense Mechanisms

By

Md. Ridwan Rahman
ID: 213-15-4435

Hemanta Roy Chowdhury
ID: 213-15-4465

FINAL YEAR DESIGN PROJECT REPORT

This Report Presented in Partial Fulfillment of the Requirements forthe **Degree of Bachelor of Science in Computer Science and Engineering**

Supervised by

Mr. Afjal Hossan Sarower
Lecturer - Senior Scale
Department of Computer Science & Engineering
Daffodil International University

Co-Supervised by

Dewan Mamun Raza
Assistant Professor
Department of Computer Science & Engineering
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY
Dhaka, Bangladesh

September 17, 2025

APPROVAL

This Project titled **"Impact of Ransomware on IoT Network Security: Analysis and Defense Mechanisms"**, submitted by Md. Ridwan Rahman, ID: 213-15-4435 And Hemanta Roy Chowdhury, ID: 213-15-4465 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **17 September, 2025.**

BOARD OF EXAMINERS

17.09.2025

Dr. Fernaz Narin Nur (FNN) — **Chairman**
Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Abdus Sattar (AS) — **Internal Examiner**
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

17.09.25

Mohammad Jahangir Alam (MJA) — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

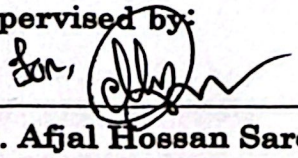
17.09.25

Dr. Sajeeb Saha (DSS) — **External Examiner**
Associate Professor
Department of Computer Science and Engineering
Jagannath University

DECLARATION

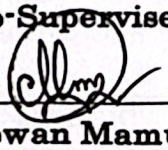
We hereby declare that this project has been done by us under the supervision of **Mr. Afjal Hossan Sarower, Lecturer - Senior Scale**, Department of Computer Science and Engineering, Daffodil International University. We also declare that neither this projectnor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:



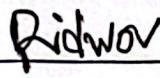
Mr. Afjal Hossan Sarower
Lecturer - Senior Scale
Department of Computer Science & Engineering
Daffodil International University

Co-Supervised by:

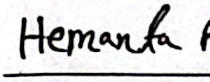


Dewan Mamun Raza
Assistant Professor
Department of Computer Science & Engineering
Daffodil International University

Submitted by:



Md. Ridwan Rahman
Student ID: 213-15-4435
Department of Computer Science & Engineering
Daffodil International University



Hemanta Roy Chowdhury
Student ID: 213-15-4465
Department of Computer Science & Engineering
Daffodil International University

ACKNOWLEDGEMENTS

This work would not have been possible without the support and contributions of many individuals over the past two semesters. We are deeply grateful to everyone who has assisted us in one way or another.

First, we express our heartfelt thanks and gratefulness to the Almighty for His divine blessing, making it possible for us to complete the **Final Year Design Project (FYDP)** successfully.

We are grateful and wish our profound indebtedness to **Mr. Afjal Hossan Sarower**, **Lecturer - Senior Scale**, Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh. Deep knowledge and keen interest of our supervisor in the field of **Network Security** to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting themat all stages have made it possible to complete this project.

We would like to express our heartfelt gratitude to the Head of the Department of Computer Science and Engineering, for his kind help in finishing our project and also to other faculty members and the staff of the Department of Computer Science and Engineering, Daffodil International University.

We would like to thank our entire course-mates at Daffodil International University, who took part in this discussion while completing the coursework.

Finally, we must acknowledge with due respect the constant support and patience of our parents.

ABSTRACT

The sudden proliferation of the Internet of Things (IoT) has brought more convenience and functionality than ever before to healthcare, education, industry, and other aspects of life; it has also created major cybersecurity risks. Of these, ransomware has become a serious menace with the ability to encrypt information or even paralyze the devices until a ransom is paid. Especially, IoT devices can be compromised because of their limited computing resources, heterogeneity, and the use of traditional security approaches that, in most cases, cannot cope with new threats. The paper will discuss the essence and effects of ransomware attacks on IoT devices, especially mobile devices and interconnected networks. We analyze ransomware dissemination processes, malware behavior, and the implications of such processes on the performance and integrity of the device. Besides, the paper assesses existing detection, monitoring, and mitigation plans, such as behavioral analysis, intrusion prevention systems, and end-user awareness programs. The study outlines the shortcomings of conventional security measures in IoT systems and highlights the need to have elaborate, customized security models. These frameworks will technically address the risk of ransomware by combining technical defenses and active user education to ensure risk reduction, safety of sensitive information, and the reliable and secure functioning of IoT ecosystems. We used nine machine learning models, and the neural network had a high accuracy of 98%. The results are critical to researchers, practitioners, and policymakers who aim to find effective processes of overcoming the ransomware threats in the networks of connected devices.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

This chapter provides the justification and background of the study discussing the rising significance of Internet of Things (IoT) in contemporary society and pertaining with it, ransomware attacks. It presents the research problem in the backdrop of current limitations of IoT infrastructures and raising demands for effective detection systems. The chapter also highlights the objectives of the research, expounds its relevance to enhancing security in IoT, and specifies the scope of work. Finally, the thesis structure (that is an elaborated content index of the following chapters) is presented.

.

1.1 Introduction

The development of the Internet of Things (IoT), which links ubiquitous devices to networks, has decisively changed the modern digital environment in its goal to bring automation and real-time data exchange to it. IoT has made impossible convenience and efficiency achievable in areas where it never existed such as in smart homes to industrial systems. However, the dynamism in cyber threats like ransomware has also brought serious security gaps. Ransomware, as opposed to traditional malware, does not only interferes with the functioning of a system, but also encrypts the essential data, which can only be restored with payment. The consequences of these attacks are further aggravated in the case of IoT ecosystems, whereby the fundamental services are breached, user security is threatened, and critical infrastructure is crippled. The nature of the IoT devices creates easy points of entry to adversaries because they have limited processing capabilities and are often not carefully monitored in terms of security settings. In turn, it is crucial to study the principles of ransomware intrusion into the framework of the IoT and determine the protective measures used against cybercrime in the context of the growing complexity of ransomware within the scope of rising and sophisticated ransomware. The proposed study aims to explore the interdependence between ransomware and IoT network security by focusing on the understanding of vulnerability, risk assessment, and the creation of practicable mitigation strategies.

Ransomware is becoming the bane of IoT networks because it takes advantage of security loopholes to cause loss of data and system downtime. It is necessary to have a clear idea of its operational dynamics and strong defensive capability [1]. IoTSDN-RAN approach proves to be more effective than the IoT SVM and IoT ML methodology in addressing the ransomware threat in the IoT system with the help of TCP packets surveillance. This highlights the need among the stakeholders to understand TCP processes to improve the security of the IoT [2]. The Internet of Things systems are becoming increasingly vulnerable to cybersecurity threats because of non-standardization and heterogeneity of devices. Aligning the countermeasures with relevant system components and incidents to boost the overall security is made possible by classifying the IoT assets based on the

attacks to which they are susceptible and their possible impacts [3]. Although wireless sensor networks (WSNs) are a part of the IoT and other fields, they face significant security issues in remote settings and limited resource situations. These deficiencies can be exploited by such well-known attacks as DoS, Sybil, wormhole, and sinkhole attacks. The paragraph outlines major threats, current protection, and future opportunities of better security provisions [4]. The increasing ransomware activity is destabilizing business organizations by encrypting and threatening to extract proprietary data and often requires high payments. Major events expose significant vulnerabilities to security and fresh strategies like double extortion [5].

This research paper aims to discuss the effect of ransomware attacks on the security of IoT networks and consider the effectiveness of countermeasures. IoT devices represent a growing collection of attackable targets by cybercriminals since they are steadily being connected to critical fields, such as smart homes and industrial systems, as well as healthcare and transportation. The main objective of this study include an understanding of the specifics of how ransomware is specifically targeted at IoT environments. The impact of these attacks on the stability of network operation, and data integrity, and the analysis of existing defensive mechanisms. Additionally, the study attempts to recommend feasible, IoT-based solutions that would boost defenses, increase the resiliency of systems, as well as aid in creating more secure and reliable IoT systems.

1.2 Motivation

The widespread adoption of Internet of Things (IoT) technologies across all critical areas, particularly in healthcare, transportation, and industrial automation, has completely changed the modern world. More operational efficiencies and real-time data sharing are made possible by such interconnected systems, but IoT devices also introduce a number of cybersecurity vulnerabilities. These days, ransomware is among the most dangerous. In contrast to ordinary malware, ransomware will not only encrypt and render private information unreadable, but it will also result in prolonged service outages and disruptions to operations. Because IoT endpoints sometimes operate on shoddy security models and lack a centralized system of protection, they are especially susceptible to these kinds of assaults. The seriousness of the scenario has been increased with the introduction of the double-extortion techniques, which involve data encryption and surreptitious exfiltration. Even though there is research in the area, few studies have addressed the dynamics of ransomware in the IoT environment. Therefore, there is an immediate need to clarify the vulnerabilities of the IoT networks and come up with strong, system-specific defense frameworks. The proposed investigation is expected to help advance resilient IoT infrastructures that can counter the emerging strategies of ransomware by discovering the current security gaps and defining the feasible attack patterns.

1.3 Objectives

The issue of security of Internet of Things (IoT) technologies has become more tense nowadays when these technologies are incorporated into the key services, such as healthcare, manufacturing, energy, and smart cities. Such networked systems are highly effective but can still be compromised by cyberattack since they are based on remote control and the real-time data exchange. Ransomware has become one of the most critical threats and transformed into a complex and devastating malware over the last few years. Ransomware, in contrast to traditional malware, not only blocks the access to important data, but also includes data exfiltration and its public release, which increases the outcome of an attack. Although the opinions about cybersecurity and IoT continue to be actively studied, the exact effects of ransomware on IoT networks are not clearly comprehended. The limited processing capabilities and memory that the IoT devices possess make most protective mechanisms cannot work, since they are typically designed to work with a more traditional computing platform. In order to bridge this gap, the current study investigates how ransomware operates in the Internet of Things ecosystem, identifies vulnerabilities, and evaluates defense strategies specific to these environments. Researchers, developers, and security experts who wish to build more robust IoT networks and make sure that ransomware attacks don't really damage society as a whole may find this study to be a useful resource. Despite the fact that IoT apps are increasingly being integrated into critical infrastructure systems, their lack of security safeguards makes them vulnerable to ransomware attacks. Because IoT networks are lightweight, even less defensive strategies can be used on them. This study fills a significant gap in knowledge by working to analyze the ransomware threats and suggesting tailored security solutions.

1.4 Methodology

This research has a methodology that is systematic in order to prepare and test an intelligent ransomware detector system to be used in an IoT environment. It starts with data collection, in which a suitable dataset of ransomware and non-malicious actions is obtained using credible sources. This dataset is processed to meet the criteria of machine learning applications such as consistency and preprocessing of the data, which involves cleaning, normalization, and encoding of the features.

This is followed by the feature selection and dimensionality reduction algorithm, e.g. normalization and standardization, which boosts model performance and minimizes complexity of computation. The resultant processed data is then trained and tested by different machine learning algorithms such as Logistic Regression, Decision Tree, Random Forest, Gradient Boosting, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Naive Bayes and Neural Networks. In a bid to deal with class imbalance, oversampling methods like the SMOTE are used.

Lastly, the standard performance metrics are used to evaluate the models in terms of accuracy, precision, recall and F1-score in order to ascertain the effectiveness of

the models to identify ransomware. A comparative analysis is done to find out the most efficient model and the findings are discussed to show the strengths and weaknesses of the proposed model.

1.5 Project Outcome

It is hoped that the expected outcome of this study is to provide a comprehensive understanding of the nature of ransomware attack and the effect of the same on the IoT settings since the reliance on the IoT networks in important sectors has been growing. One of the key outcomes is to find the exact vulnerabilities of the IoT systems that attract cybercriminals to them. The study also aims at analyzing the different types of ransomware, their modes of transmission, and how they exploit the vulnerability of IoT devices by compromising their security features. A comparative analysis of the security measures and defenses already in place to stop or manage ransomware in IoT networks is another anticipated result. It is anticipated that the research would identify the shortcomings of current solutions and suggest modifications by looking at recent examples and real-world case studies. Proposing some workable, resource-efficient defense tactics that are especially compatible with the limitations of IoT infrastructure will be one of this study's contributions. A conceptual framework or model that enables companies to proactively monitor, identify, and mitigate ransomware risks in IoT systems is another anticipated outcome of the study. The finished product aims to support developers, IT experts, and researchers by providing real-world knowledge and recommendations on how to reinforce the IoT security against the threat of ransomware attacks that are only growing stronger.

1.6 Organization of the Report

Six chapters make up my research study: Introduction, Background, Research Methodology, Experimental Results and Discussion, Impact on Sustainability, Environment, and Society, and, lastly, Summary, Conclusion, Recommendation, and Implications for Further Research.

Chapter 1 presents the idea of ransomware threats in Internet of Things contexts, laying the groundwork for why this subject is so important in the connected world of today. The second chapter examines the body of research on ransomware attacks and IoT vulnerabilities and identifies the main obstacles to IoT network security. Chapter 3 presents the suggested protection methodology, which includes comprehensive flowcharts and system architecture. It describes the multi-layered strategy intended to identify and lessen ransomware threats. Experimental results are presented in Chapter 4, which assesses how well different machine learning and anomaly detection methods recognize ransomware behaviors. To highlight each model's advantages and disadvantages, a comparative study is provided. While evaluating the sustainability and long-term viability of the suggested security systems, Chapter 5 examines the wider effects of ransomware assaults on the environment and society. The last chapter highlights the importance of the study, main conclusions, and suggests possible future paths for improving IoT security against changing ransomware attacks.

Chapter 2

Background

This chapter emphasizes the fast development of IoT and its increasing security issues, especially ransomware attacks which use the vulnerabilities of their devices and the lack of their security. In this chapter also analyses the literature that suggests solutions including intrusion detection, lightweight cryptography, blockchain, and AI-based ones. The literature highlights the necessity to develop adaptable and scalable security mechanisms in IoT settings.

2.1 Introduction

The explosive nature of the Internet of Things (IoT) has totally changed device communication, data collection, and sharing and IoT networks have become deeply embedded in daily life, essential infrastructure and other things. It includes smart homes, industrial automation, smart agriculture and so on. However, there exist some serious security concerns due to the growth of massive IoT devices. One of the largest threats is ransomware, which is a malicious software similar to encryption programs but requiring payment to unlock the data that was encrypted. In IoT networks, where real-time operations are a must. Ransomware can disrupt the operations severely, compromise the privacy of users and even pose a threat to physical safety. To defend against ransomware and other IoT device attacks, one needs to know their nature, their behavior within the network environment, and how it propagates across the IoT environment. The IoT devices are often poorly secured, and this factor predisposes them to being targeted by attackers in comparison to conventional systems. Moreover, the heterogeneity of the devices and the low processing power makes the process of adopting conventional security solutions more complex. This section of the study provides a summary of the key concepts related to ransomware attacks, network security, and Internet of Things. Also, it discusses the underlying technologies and communication schemes such as MQTT, CoAP, and 6LoWPAN and other security things and communication protocols, which are used in IoT systems. It goes further to explain the general weaknesses that exist in IoT-based devices and the underlying attack vectors that ransomware uses or attackers use to steal the data. By forming these preliminary conclusions, this research forms the basis of a more detailed investigation into the impact of ransomware and the development of advanced protection mechanisms that could be used on IoT networks.

2.2 Literature Review

The study highlights the rapid growth of IoT gadgets and devices used in daily life that have created risks of ransomware assaults due to low security, minimal upgrading, and poor authentication. Several studies have shown that the risk can be demonstrated by events like the Mirai botnet. Researchers have suggested some ideas for solutions such

as blockchain, intrusion detection systems based on machine learning, and lightweight encryption. But still there is a lack of a global standard of IoT security. Moreover, to prevent ransomware attacks going forward specialists emphasize the urgent need to have design-level security in IoT equipment. [1].

Recent research also indicates that ransomware infiltration of Internet of Things devices is rapidly on the rise and that has escalated the demands of the state-of-the-art security tools. Various tests have shown that the IoTSDN-RAN concept is an effective and profitable solution that reduces the ransomware threat significantly. IoTSDN-RAN is also more reliable than IoT SVM due to the fact that, based on comparative research, the False Negative Rate (FNR) of IoTSDN-RAN is lower. Our approach that uses SDN has improved ransomware detection in the context of IoT. Based on the results, preventative measures may be even more effective considering the risks of their occurrence than the prevention of those threats through detection and mitigation. [2].

The research by Hafeez et al. (2020) will give an in-depth overview of cyberattacks on the IoT systems by mapping them against assaults to the IoT architectural layers, assets, and threat surfaces. This paper proposes a stratified taxonomy of attacks (physical, network, and application layer) along with their unique impact on the availability, integrity, and secrecy, as opposed to a significant number of earlier sources, which focus on individual threats. The authors are clear about the threat response mapping as they present 36 different types of attacks and 50 suitable remedies. Work highlights the absence of standardization in IoT security structures and the necessity of real-time detection in the work as one of the key contributions. This literature forms the basis of developing a set of all-encompassing and active IoT security. [3].

Wireless Sensor Networks (WSNs) are vulnerable because of their deployment in hostile environments and limited computational resources, according to recent studies. To counteract DoS, Sybil, wormhole, and sinkhole attacks, researchers have suggested a variety of protection strategies, including key distribution, encryption, and intrusion detection systems. For increased privacy, Tin Yang suggested a three-factor authentication protocol, while Zhang et al. introduced self-organizing key management. Others concentrated on time synchronization problems, attack detection algorithms such as NB-DPC and SF-APIT, and risks unique to the healthcare industry. WSNs in particular contexts, such as underwater and vehicular networks, are also examined in the literature. The most common threats in WSNs are DoS (30%) and sinkhole (23%), which are followed by Sybil (19%) and spoofing-data tampering (15%) attacks. All things considered, current research focuses on low-resource, adaptive security solutions that are adapted to WSN restrictions. [4]

Ransomware has quickly developed into a serious cybersecurity risk by taking advantage of both human and technological weaknesses. Real-world ransomware occurrences in industries including government, healthcare, and education are examined by Temara (2023). According to the study, breaches are facilitated by antiquated systems, ignorance, and inadequate authentication. The use of phishing and exploit kits in attack deployment, as well as encryption-based extortion, have been highlighted in earlier

studies. The study compiles information from occurrences around the world, including insights into the prevalent vectors, recovery difficulties, and attacker behavior. This body of literature emphasizes how urgently proactive defenses and cybersecurity training for staff members are needed. [5]

The dynamic character of ransomware has been a subject of deep investigation during the last several years, with a specific focus on its growing complexity and catastrophic impact on the protection of data. Through analyzing their mechanisms of operation and possible points of entry such as downloading of malware and phishing emails, researchers have categorized ransomware into types such as crypto-ransomware and locker-ransomware. Many research papers emphasize on the complete lifecycle of an attack, such as ransom demands, encryption, and intrusion. Attention has been paid to the prevention and mitigation methods, including behavioral detection, regular backups, and robust security architecture. The literature review also addresses the psychological tricks employed to make the victims pay. All said and done, these results emphasize the need to have and act proactively and flexibly in ensuring cybersecurity for IoT devices and data. [6].

The number of contributions to literature also grew tremendously in the period between 2019 and 2024, that is an indication of the rising importance of credible defense strategies in the context of AI use in security practices. This literature review of the study underscores the acuity with which interdisciplinary strategies must be adopted with a view to effectively combating emerging aggressive dangers. [7].

To counter such risks, many scholars have explored various methods, such as anomaly detection, intrusion detection systems (IDS), adversarial training, and cryptographic protections. The recent IoT security trend shows that the number of papers from 2019 to 2024 is significantly higher, most of them indicating a growing focus on securing AI systems against evolving adversarial threats in academia and the industry at large. Such a shifting trend of inquiry highlights the necessity of smart, adaptable and scalable defense infrastructure for IoT devices. [8].

The literature lists dangers that take advantage of the diverse and resource-constrained environment of the Internet of Things, including ransomware, botnets, denial-of-service, man-in-the-middle, and social engineering. Scholars draw attention to the shortcomings of conventional security methods and stress the necessity of decentralized, scalable, and lightweight solutions like blockchain. Securing IoT infrastructure continues to be a major research and implementation challenge as its use grows quickly. [9]

The concerning increase in ransomware attacks directed at Internet of Things devices is fueled by the quick development of smart technology. IoT networks are particularly vulnerable, according to researchers, because of their diverse topologies, low processing power, and inadequate built-in security features. Numerous studies examine new attack methods, including DDoS, botnets, and man-in-the-middle attacks, and suggest security tactics like intrusion detection, lightweight cryptography, and blockchain-based frameworks. There is a need for more robust and context-aware security models since,

despite increased awareness, existing solutions frequently lack scalability and adaptability. [10]

Since Android and IoT systems have open architectures and are widely used, numerous studies have shown how quickly malware threats are expanding to target these platforms. Focusing on the behavior and infection vectors of ransomware, malware, and trojans, researchers have examined a variety of attack types. The effectiveness of current defense methods, such as machine learning models and signature-based detection, has been demonstrated to be imperfect. Limitations in scalability and flexibility to changing threats are still an issue, though. The necessity of hybrid models that combine static and dynamic analysis for improved threat mitigation has been emphasized in recent publications. [11]

Recent studies demonstrate the growing complexity of ransomware attacks and their expanding influence on the workplace, healthcare, and critical infrastructure settings. While earlier research concentrated on simple encryption-based attacks, more recent studies highlight strategies like double and triple extortion and the rise of ransomware-as-a-service (RaaS). The shortcomings of antiquated backup plans and conventional signature-based defenses are also mentioned in the literature. Adaptive solutions like blockchain for data integrity, AI-driven threat detection, and Zero Trust architectures are being promoted by academics. A change from reactive to proactive security methods that can adapt to ransomware threats is required by this growing body of work. Building on these discoveries, the current review investigates novel mitigation strategies. [12]

Malware threats in Android and IoT systems are always changing, taking advantage of obsolete firmware, insecure communication channels, and weak authentication. Ransomware, trojans, worms, and spyware are examples of common attack types that jeopardize device operation and user privacy. Although they provide some protection, defensive techniques like sandboxing, antivirus software, and permission control models are ineffective against zero-day attacks. To enhance threat identification, hybrid detection systems, behavioral analysis, and machine learning techniques are being developed. Reducing vulnerability to these attacks requires robust access control and firmware-level security. To handle the increasing complexity of these threats, a multi-layered and flexible defense approach is necessary. [13]

Several security challenges are addressed by wireless sensor networks (WSNs) because of the limited resource base and the deployment in vulnerable settings of IoT devices. Attacks such as DoS, Sybil, sinkhole, and wormhole are common and attack a variety of layers on the network layer. To overcome these challenges and secure the IoT devices, researchers have come up with techniques, such as lightweight cryptographic protocols, intrusion detection systems, and key distribution. Also, the approaches vary in health care, as well as in the military and in other fields of IoT. The considered research works demonstrate the urgency of effective and scalable solutions for IoT devices and data that do not deteriorate the performance of the network without exhausting the network resources. [14].

Some of the extreme security threats that the Wireless Sensor Networks (WSN) have to deal with include Sybil, sinkhole, wormhole, and Denial-of-Service (DoS) attacks severely. These attacks undermine data confidentiality, routing processes, as well as network structure. Researchers have studied various countermeasures such as intrusion detection systems, safe routing protocols and cryptographic algorithms to keep IoT environment safe and sound. Distribution of keys, clustering and authentication systems with two-factor authentication have also been proven to be effective in enhancing security. However, the creation of scalable, energy-efficient, and lightweight security solutions that are applicable in the WSN environment with low resources still presents challenges. [15].

Table 2.1: Summary of Literature Reviewed.

Author(s)	**Year**	**Title**	Methodology	**Key Findings**
Humayun et al. [1]	2020	Internet of Things and Ransomware: Evolution, Mitigation and Prevention	Conceptual review and synthesis of prior research	Rapid IoT adoption increases ransomware risks due to weak security and poor updates; proactive design-level security is critical.
Khan & Rahman [2]	2023	Efficiency of Surveillance of TCP Packets in IoT in Reducing the Risk of Ransomware Attacks	Simulation-based experimental analysis of TCP packet monitoring	Surveillance of TCP packets in IoT significantly reduces ransomware risk; outperforms IoT SVM with lower False Negative Rate.
Msgna [3]	2022	Anatomy of Attacks on IoT Systems: Review of Attacks, Impacts and Countermeasures	Structured literature survey with taxonomy building	Provides a layered taxonomy of IoT attacks; identifies 36 attacks and 50 mitigation strategies; highlights lack of standardization.
Alotaibi et al. [4]	2025	Assessment of Cybersecurity Threats and Defense Mechanisms in Wireless Sensor Networks	Comparative study of existing security models in WSNs	Reviews recent studies on enhancing WSN security; discusses challenges and proposes defense mechanisms.
Temara [5]	2024	The Ransomware Epidemic: Recent Cybersecurity Incidents Demystified	Case-based analysis of real-world ransomware events	Human and technological weaknesses contribute to ransomware spread; emphasizes proactive defense and staff cybersecurity training.

Reshmi [6]	2021	Information Security Breaches Due to Ransomware Attacks – A Systematic Literature Review	Systematic mapping of peer-reviewed ransomware studies	Examines various types of ransomware, vulnerabilities, attack methodologies, impacts, and mitigation techniques.
Khaleel et al. [7]	2024	Network and Cybersecurity Applications of Defense in Adversarial Attacks: A State-of-the-Art Using ML and DL	Survey of ML/DL defense techniques in adversarial settings	Investigates defense strategies in adversarial attacks using ML and DL methods; synthesizes existing research.
Jimmy [8]	2024	Cyber Security Vulnerabilities and Remediation Through Cloud Security Tools	Exploratory literature-based assessment	Surveys cyber attack patterns and detection methodologies; discusses remediation through cloud security tools.
Anjum et al. [9]	2021	Analysis of Security Threats, Attacks in the Internet of Things	Comprehensive literature review	Analyzes IoT security threats and attacks; discusses mitigation strategies and challenges.
Mousa & Shehab [10]	2025	Applying Risk Analysis for Determining Threats and Countermeasures in Workstation Domain	Risk assessment framework application	Applies risk analysis to determine threats and countermeasures in workstation domains.
Yadav et al. [11]	2022	Malware Analysis in IoT & Android Systems with Defensive Mechanism	Survey and categorization of malware detection methods	Reviews malware threats in IoT and Android systems; discusses defensive mechanisms.
Ogini et al. [12]	2025	Impact of Ransomware Attack Trends on the Effectiveness of Defense Mechanisms	Narrative review with trend-based evaluation	Reviews ransomware attack trends; discusses the effectiveness of current defense mechanisms.
Costin & Zaddach [13]	2018	IoT Malware: Comprehensive Survey, Analysis Framework and Case Studies	Survey and case study approach	Provides a comprehensive survey of IoT malware; presents an analysis framework and case studies.
Saeed et al. [15]	2020	Ransomware: A Framework for Security Challenges in Internet of Things	Framework design and conceptual validation	Develops a framework addressing security challenges posed by ransomware in IoT environments.

The reviewed studies highlight the rising threat of ransomware in IoT systems, using approaches such as literature surveys, case analyses, experimental models, and framework design. Findings show that traditional defenses are inadequate, with vulnerabilities linked to weak updates, poor standardization, and human factors. Most works stress the need for lightweight cryptography, AI/ML-driven solutions, and proactive security frameworks to safeguard IoT environments.

2.3 Gap Analysis

Several additional studies have explored IoT security from complementary perspectives that are not captured in the main comparison table. Ashrafuzzaman et al. [19] analyzed IoT cybersecurity threats using leaked credentials and honeypots, emphasizing the risks posed by exposed device credentials, which most IDS approaches overlook. Sharma and Bhushan [22] proposed federated learning-based intrusion detection frameworks, but their work primarily focused on algorithmic performance without integrating ransomware-specific analysis or explainable AI. Such and Kamel [23] surveyed adversarial robustness in network intrusion detection, highlighting vulnerabilities that could compromise IoT IDS models under adversarial attacks, a gap not fully addressed in existing systems. Qiu et al. [24] studied adversarial intrusion detection using the NF-ToN-IoT-v2 dataset, pointing out the limitations of standard datasets in evaluating model resilience to crafted attacks. Kim et al. [27] investigated explainable AI for IoMT security using XAI-XGBoost, demonstrating improved model interpretability, which remains underutilized in current IoT IDS frameworks. Incorporating insights from these studies can strengthen the proposed research by addressing adversarial threats, credential-based vulnerabilities, and explainable AI integration alongside ransomware-specific detection and lightweight defense mechanisms [19], [22], [23], [24], [27].

Table 2.3.1 Gap Analysis of existing work and proposed study

Feature / Focus Area	M. Ashrafuzzaman et al. [16]	A. Dabbagh et al. [17]	Y. Inoue et al. [18]	H. E. Al-Rakhami et al. [20]	N. C. Luong et al. [21]	A. Souri et al. [25]	A. Abdelaziz et al. [26]	P. Verma et al. [28]	R. K. Singh et al. [29]	Proposed Research
Ransomware-specific analysis	No	No	Yes	No	No	No	No	No	No	Yes
IoT/WSN attack taxonomy	No	No	No	No	No	No	No	No	No	Extended
Use of real-world honeypot data	No	Yes	Yes	No	No	No	No	No	No	Yes
Experimental validation with IoT datasets	Yes	No	No	Yes	Yes	No	No	No	No	Yes
Lightweight/energy-efficient defense	No	No	No	No	No	Yes	No	No	No	No
Adversarial robustness	Nos	No	No	No	No	No	Yes	No	Yes	Yes
Human/organizational factor focus	No	No	No	No	No	No	No	No	No	Yes
Compliance/standardization issues	No	No	No	No	No	No	No	No	No	Yes

In the proposed research, all key focus areas are addressed to bridge existing gaps in IoT security. This includes ransomware-specific analysis, an extended IoT/WSN attack taxonomy, and the use of real-world honeypot data for realistic threat modeling. The study incorporates experimental validation with IoT datasets collected from Kaggle, lightweight and energy-efficient defense strategies, and mechanisms to ensure adversarial robustness. Additionally, human and organizational factors are considered, alongside compliance with standardization requirements, making the framework comprehensive and practical for real-world IoT environments.

2.4 Summary

The chapter examined the existing body of knowledge in the field of IoT security, ransomware, and intrusion detection system. The current literature addresses ransomware detection, IoT/WSN attacks taxonomy, federated learning, lightweight defense strategy, and explainable AI integration. Nevertheless, these aspects are covered in most works partially or as an individual aspect, with disjointed IoT security solutions. The limitations of the previous research mentioned in the gap analysis table are a lack of ransomware specificity, the lack of usage of real-life data, experimental validation, as well as insufficient attention to human and organizational factors. To mitigate these shortcomings, the proposed study integrates ransomware-specific study, a comprehensive list of IoT/WSN attacks, real-world honeypot measurements, lightweight and energy-on-demand defense measures, federated learning, adversarial resilience, explanation-based AI, and compliance with standardization and human factor aspects. This integrative methodology would be trying to come up with a stronger, practical and holistic model of securing the IoT environments.

Chapter 3

Research Methodology

This chapter gives the research methodology to be used in this study. It explains how to prepare the dataset, preprocess it, and address the problem of class imbalance by oversampling. The chapter also describes the systematic processes involved to make the dataset clean, balanced, and to analysis with the help of machine learning.

3.1 Methodology

The research methodology will be used to constitute an organized way of designing a successful ransomware detection system in IoT settings and environments. It provides a chronological procedure, beginning with the collection and preprocessing of the datasets, all the way to the class imbalance and the data arrangement before training the machine learning models. All the stages guarantee that the information is purged, harmonized, and this is amenable to valid analysis and assessment.

3.1.1 Overview

The methodology considered in this research focuses on preparing and balancing the IoT security dataset to ensure reliability in subsequent machine learning analysis. To prevent IoT attacks, the process begins with the identification of relevant requirements. It includes the selection of features and the definition of the target variable. After that, data preprocessing techniques are applied. The preprocessing are such as label encoding for categorical identifiers, one-hot encoding for categorical attributes, and standardization of numerical features to maintain uniformity across variables.

To address the issue of class imbalance in the dataset, two complementary oversampling strategies are employed. One is the Synthetic Minority Oversampling Technique (SMOTE) to generate synthetic samples of the minority class. Another is Random Oversampling (ROS) to increase the representation of the majority class. These two combined approaches result in a balanced dataset with equal class distribution of the target class. Finally, the preprocessed and augmented dataset is reconstructed with appropriate feature names. Then the dataset is stored for further experimentation and evaluation in this study.

3.1.2 Proposed Methodology

The approach used in this study is the preparation and balancing of the security dataset of the IoT, so that the latter becomes reliable when subsequently subjected to machine learning. The experiment starts with the identification of the pertinent requirements, such as features selection and the definition of the target variable. The process of data preprocessing is performed including one-hot encoding of categorical identifiers, label

encoding of categorical attributes, and normalization of numerical values. Those pre-processing ensure that the variables are the same.

To overcome the problem of class imbalance in the data, two complementary oversampling methods are used. Synthetic Minority Oversampling Technique (SMOTE) is used to create synthetic examples of the minority class. Oversampling (ROS) is applied to boost the representation of the majority class. The result of this two preprocessing and combination methods generates a balanced dataset with equal distribution in classes. The last step in the data preprocessing is the reconstruction of the preprocessed and enhanced dataset with the correct feature names and is left to be harnessed in further experiments and analysis.

I. Requirement Analysis

It starts with an analysis of requirements that entails a clear articulation of the objectives of the study. The features that are identified at this stage are presented and the target variable (Threat Mitigated) is chosen as the most important outcome measure. The dataset is also analyzed in terms of completeness, consistency, and general appropriateness to preprocessing. This will allow missing values, inconsistencies, or irrelevant fields to be prevented in the way of subsequent steps.

II. Data Collection and Preparation

Once the requirements are established, the dataset is imported from its source for preprocessing. Since raw data often contains categorical and numerical attributes in mixed formats, suitable transformations are applied. Identifiers such as *Device ID* and the binary target variable are transformed using label encoding, converting them into numeric form without introducing artificial complexity. Nominal features such as *IoT Layer*, *Request Type*, *Security Threat Type*, and *Consensus Mechanism* are transformed using one-hot encoding, which represents each category as a distinct binary column, thereby preventing false ordinal relationships. Furthermore, numerical features are standardized using Z-score normalization, which scales the values to have a mean of zero and unit variance. This standardization step ensures that all features contribute equally during model training, regardless of their original scale.

III. Class Imbalance Handling

A preliminary analysis of the dataset reveals a significant imbalance between the classes of the target variable. In imbalanced datasets, machine learning models tend to be biased toward the majority class, resulting in poor predictive performance for the minority class. To address this, the distribution of the target variable is analyzed, and resampling techniques are planned to balance the dataset. Importantly, resampling is applied only to the training set, leaving the test set untouched, to prevent data leakage and preserve the reliability of model evaluation.

IV. Oversampling using SMOTE and ROS

Two complementary oversampling methods are used to fix the imbalance in the classes. The Synthetic Minority Oversampling Technique (SMOTE) is used in the first step, which is the production of synthetic samples of the minority. Contrary to the random duplication, SMOTE generates new and realistic examples through interpolating among existing minority ones, which enhances the generalization. After the size of the minority class is reached, Rand Dominating Oversampling (ROS) is employed to the majority class to balance the distribution even more and to have both classes represented equally. This hybrid approach is a combination of the benefits of synthetic data generation and straightforward duplication, which leads to a balanced dataset of large enough size to train on.

V. Dataset Reconstruction and Export

After all the preprocessing and oversampling, the dataset is reassembled to a structured format. This format can be used for machine learning. All names of the features are reattached transparently. After that, the target variable (Threat Mitigated) is added as the result column. After that, the dataset is stored in a CSV format. The final balanced data is now prepared to train the model, test it and carry out additional experiments. This guarantees that the dataset is well-documented to ensure reproducibility as well as being balanced and standardized for better training of machine learning models.

The methodology follows some well-planned steps; the datasets about internet security available are converted to a clean, consistent, and balanced form. The proposed workflow, tackling the preprocessing and imbalance issues, that creates a strong base for the following predictive modelling and experimental analysis.

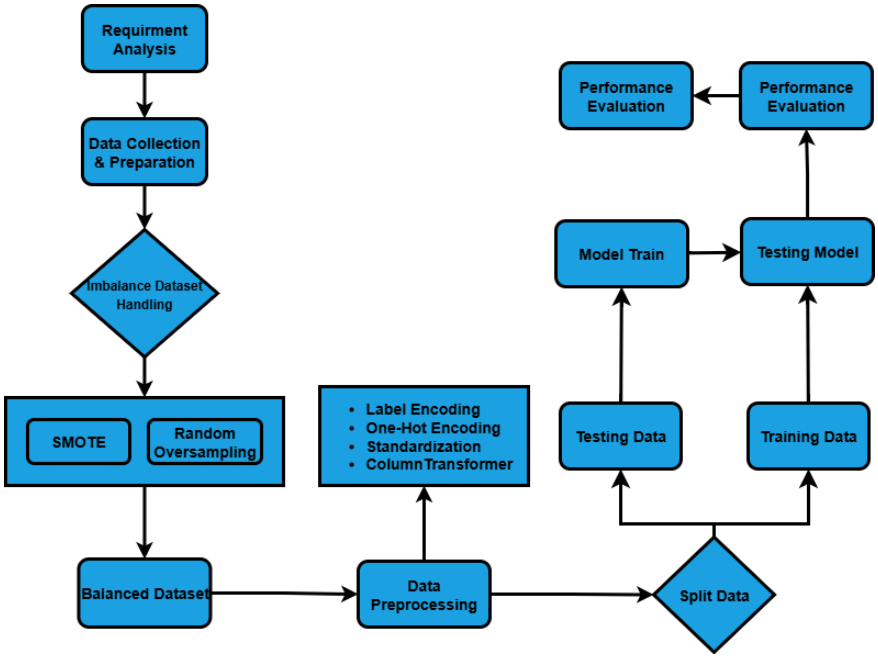


Figure 3.1.2.1: Proposed Methodology of the research

The suggested methodology offers a systematic and elaborate approach to an IoT security dataset preparation to analyze the threat with the help of machine learning. It starts with a strict requirement analysis, which determines the objectives of the study, the identification of useful features, and the completeness and consistency of the dataset, thus, providing the basis of credible preprocessing. The next step in the data collection/data preparation converts all categorical variables into a numeric form via label and one-hot encoding, and standardizes numerical features via Z-score normalization, ensuring that all features contribute equally to training the model. It is known that the imbalance in the classes is a challenge, and, therefore, the methodology utilizes specific resampling methods to ensure that the model is not biased towards the majority class. The combination of SMOTE with Random Oversampling is effective in balancing the classes, which produces artificial minority samples and reduces the number of majority classes, which improves the generalization of models. Lastly, the dataset is finally reconstructed as a well-documented structure with clear feature names with an appended target variable and exported in CSV format so as to enable reproducibility and future experimentation. Generally this multi-stage methodology will provide clean, standardized, and balanced data, which will serve as a strong base to predictive modelling and precise threat detection in IoT settings.

3.2 Detailed Methodology and Design

The methodology to be used in this study will seek to convert the raw dataset of the IoT security into a structured, standardized and balanced data, which can be used in machine learning to predict threats. All the steps of the methodology were chosen based on the analysis of the options presented to guarantee the most appropriate choice of data integrity, model readiness, and performance. The steps that are followed are as detailed below..

3.2.1 Dataset Preparation

The data was first gathered on Kaggle, which has both categorical and numeric variables, as well as binary target classification: Threat Mitigated. The first step was the cleaning of the data through the process of processing missing values, fixing inconsistencies and standardizing of column names. One-hot encoding was used to encode categorical features i.e. IoT Layer, Request Type, Security Threat Type and Consensus Mechanism, whereas standardization was used to scale numerical features i.e. Data Size, Processing Time, and Energy Consumption to reach similar ranges. Also, the data was checked on the presence of class imbalance, which was subsequently corrected with the help of SMOTE and Random Oversampling to create a balanced dataset to be used in threat analysis, based on machine learning.

Device ID	IoT Layer	Request Type	Data Size (KB)	Processing Time (ms)	Security Threat Type	Attack Severity (0-10)	Blockchain Transaction Time (ms)	Consensus Mechanism	Energy Consumption (mJ)	Threat Mitigated
D0001	Application	Data Transmission	552	42	DDoS	1	142	PBFT	1.26	1
D0002	Application	Data Transmission	704	24	DDoS	1	197	PoS	0.66	0
D0003	Device	Authentication	485	35	Eavesdropping	7	272	PoS	0.92	0
D0004	Application	Data Transmission	965	21	Eavesdropping	1	187	PoS	1.88	0
D0005	Device	Encrypted Data Transfer	653	17	Tampering	1	155	PoS	1.8	1
D0006	Application	Smart Contract Execution	265	9	Tampering	2	112	PoS	1.81	0
D0007	Application	Data Transmission	425	24	Tampering	5	252	PoA	1.49	1
D0008	Application	Authentication	940	19	Eavesdropping	9	230	PoS	1.47	1
D0009	Network	Encrypted Data Transfer	1337	15	Eavesdropping	4	201	PBFT	1.1	1
D0010	Application	Smart Contract Execution	704	42	Eavesdropping	10	159	PoS	1.89	1

Fig 3.2.1.1 Dataset Overview

Threat Mitigated target variable in the dataset is coded into 2 categories 0 and 1. The Class 0, which consists of threats not addressed, has 717 samples that constitute 71.7 percent of the overall data. Class 1, in turn, that is the threats that were neutralized, contains 283 samples, which occupy 28.3 percent of the data. This apparent imbalance indicates that most of the cases fall under Class 0 whereas Class 1 is underrepresented. This imbalance can lead to machine learning models being biased to predict the majority class, which demonstrates the need to use balancing algorithms such as Random Oversampling (ROS) and SMOTE to achieve fair and reliable models training.

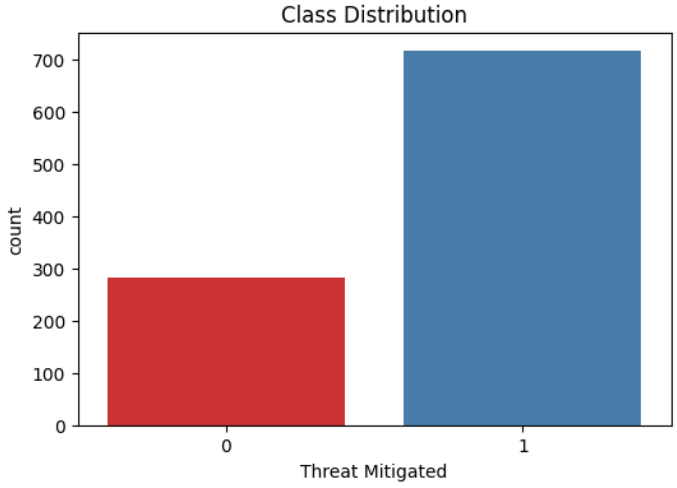


Fig 3.2.1.2 Class Distribution of Threat Mitigated

Class Distribution of Categorical Features: The dataset contains several categorical features that capture different aspects of the IoT system and its security environment. These include Device ID, which uniquely identifies each IoT device; IoT Layer, which specifies the network layer (e.g., perception, network, or application) where the data or attack originates; Request Type, which indicates the type of request being made within the system; Security Threat Type, which categorizes the nature of the security threat (e.g., denial-of-service, spoofing, or unauthorized access); and Consensus Mechanism, which refers to the blockchain-based validation protocol employed (e.g., PoW, PoS, PBFT).

Since these features are non-numeric, preprocessing was required before feeding them into machine learning models. Specifically, Device ID was transformed using label encoding, while the other categorical features were converted using one-hot encoding. This ensured that categorical variables were properly represented in numerical form without introducing ordinal bias. These features are crucial because they provide contextual information about devices, attack types, and blockchain mechanisms that directly influence the detection and mitigation of threats in the dataset.

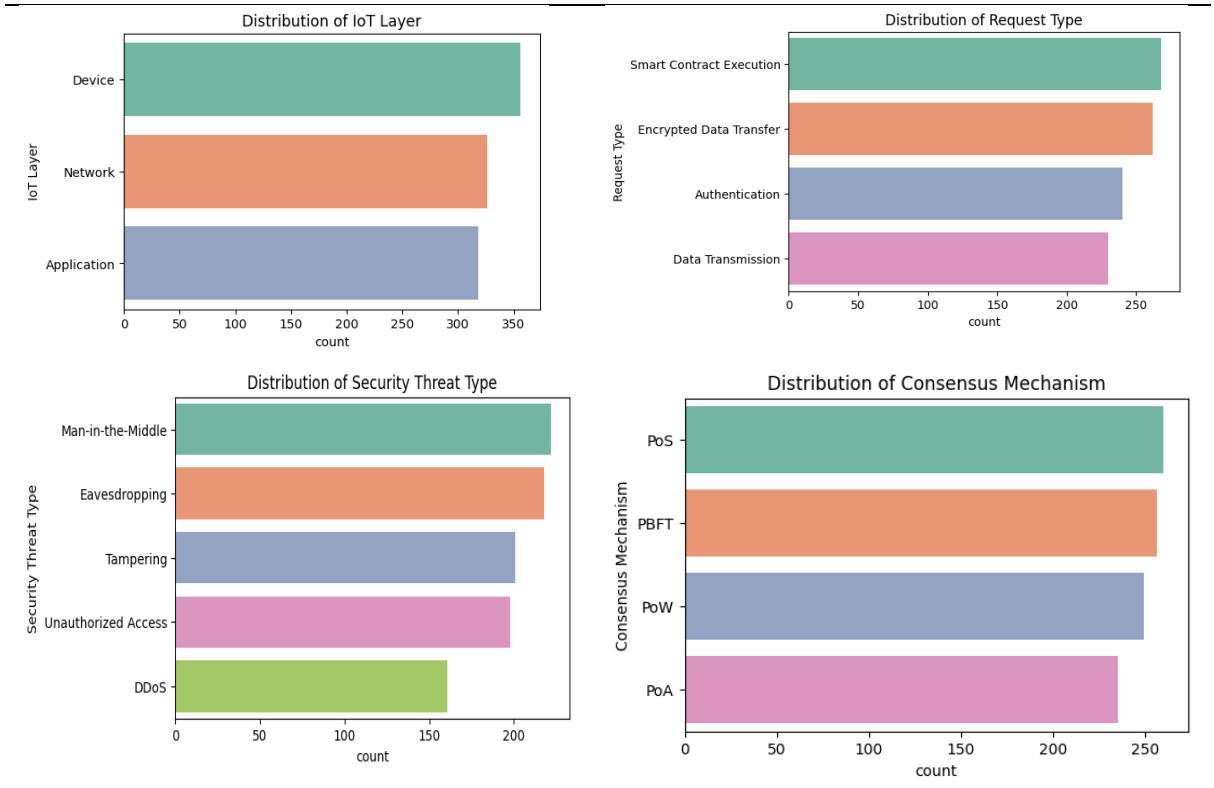


Fig 3.2.1.3 Distributions of Categorical Features

Class Distribution of Numerical Features: The dataset includes a range of numerical characteristics that give the most important quantitative data regarding the work of the IoT systems and the nature of security threats. These characteristics are continuous variables that ease modeling, analysis, as well as prediction of threat mitigation results. The most important numerical characteristics are as follows:

I. **Data Size (KB):** The feature is a measure of the data that is exchanged or handled by the IoT devices. The size of the data can become heavier to the network and cause higher processing delays that can impact the performance of the system and make some attacks more effective. The size of data analysis assists in determining the relationship between the consumption of resources and the mitigation of the threats.

II. **Processing Time (ms):** It is an attribute that quantifies how many milliseconds it takes to run a request or operation in the IoT system. Long processing times is a sign of bottlenecks, inefficient algorithms or active attacks. Incorporation of this variable enables the model to be responsive to the system in terms of mitigating a threat.

III. **Attack Severity (0–10):** This is a scaled data that determines the attack severity or the intensity of the detected attack with high values representing severe attacks. The inclusion of the severity of the attacks gives the model context on the possible effects of a threat which is important in prioritizing mitigation measures.

IV. **Blockchain Transaction Time (ms):** This parameter can be estimated in systems which operate blockchain as a part of an IoT security framework to log or validate transactions in blockchain. The consensus mechanism and network load determine transaction time, and threats to real-time are influenced by it. This feature can be used to model and estimate a trade-off between security and performance.

V. **Energy Consumption (mJ):** Energy Consumptionfeature measures the energy that is used to carry out a certain operation or transaction. IoT devices tend to be energy limited, and high energy usage can imply inefficiency or heavy operations.

All numerical variables were preprocessed in Z-score standardization in order to make them fit in machine learning models. The data in this technique are transformed in a way that every feature has zero mean and standard deviation of one, and the features with big scales do not override the ones with smaller scales. Algorithms that are sensitive to the magnitude of features, including Support Vector Machines, k- Nearest Neighbors, and Neural Networks are especially sensitive to standardization.

A comparison in terms of statistical summaries (mean, minimum, maximum and standard deviation) of these numerical characteristics was also investigated to find out possible outliers, skewed distributions and variation within the dataset. This discussion can guide feature engineering and can be used to design robust and balanced threat mitigation models.

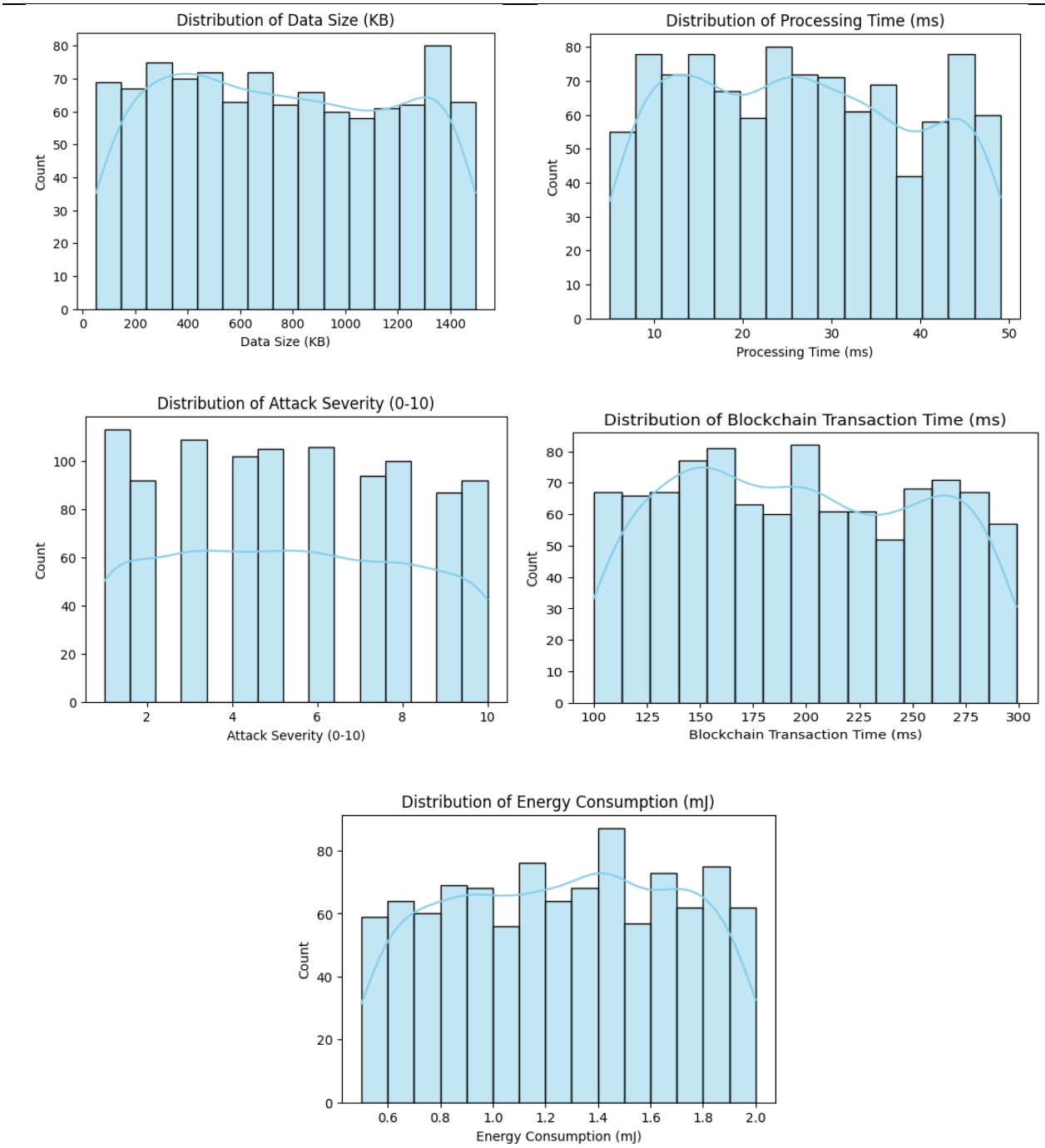


Fig 3.2.1.4 Distributions of Numerical Features

Scatterplot Matrix (Pair plot):

A pair plot also called a scatterplot matrix was drawn to examine the connections between numerical variables and their relationship with the target variable, Threat Mitigated. This visualization will be used to plot each numerical feature against all others in a grid format, which will give an overall picture of interaction of features with each other.

- The distribution of the individual features are plotted (usually as histograms or kernel density estimates) on the diagonal plots, indicating the distribution of features, their skewness, and whether they have outliers or not.

- The off-diagonal plots present the pairwise scatterplots that aid the discovery of both linear and non-linear associations, clusters, and likely patterns that distinctly differentiate the classes 0 (Not Mitigated) and 1 (Mitigated).

The pair plot allows separation of the two classes using a hue that depends on a target variable to display the two classes in different colours. It is easy to detect the ability of numerical characteristics to differentiate successfully between mitigated and non-mitigated threats. As an example, such characteristics as Attack Severity or Processing Time can represent different clusters in each of the classes, which suggests that they can be predicted.

All in all, the scatterplot matrix is an extensive exploratory data analysis (EDA) instrument. It provides direction to the selection of features, assists in the identification of multicollinearity and enlightens the model construction approach as it helps to identify the most discriminative features on the target outcome.

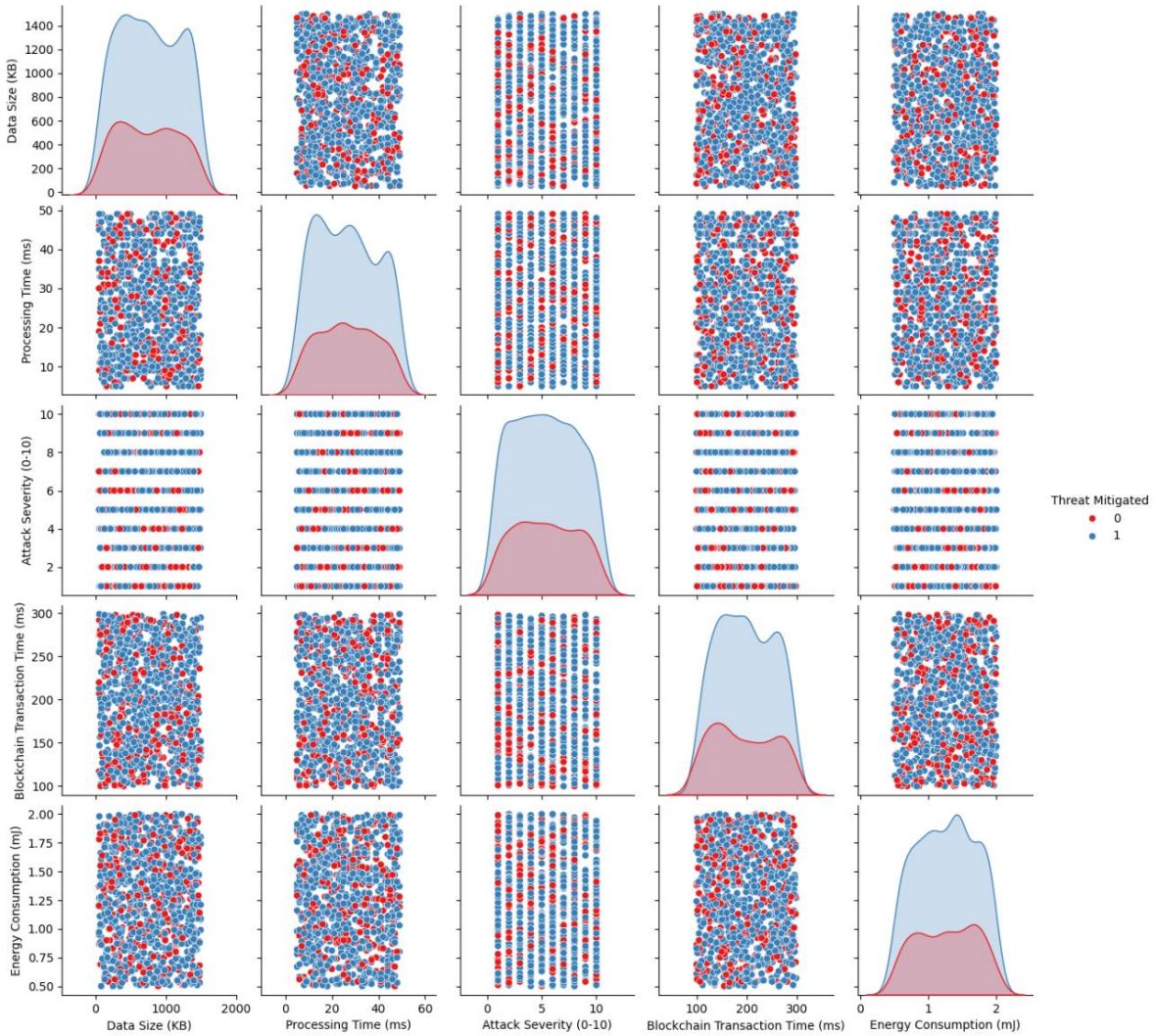


Fig 3.2.1.5 Scatterplot matrix (Pair plot)

Correlation Matrix

To explore the linear relationships between the numerical features and how they are related to the target variable Threat Mitigated a correlation matrix was produced. In this matrix, cell means the Pearson correlation coefficient of two variables and values will be between -1 and 1:

- +1 indicates a perfect positive correlation
- -1 indicates a perfect negative correlation
- 0 indicates no linear correlation

The correlation matrix provides several insights:

1. **Feature relationships:** By observing correlations among features such as *Data Size*, *Processing Time*, *Attack Severity*, *Blockchain Transaction Time*, and *Energy Consumption*, it is possible to identify highly correlated variables that may carry redundant information. This is useful for dimensionality reduction or feature selection.

2. **Target association:** The correlation of each numerical feature with *Threat Mitigated* helps identify which features are most predictive of the target class. Features with stronger correlations can be prioritized during model development.

3. **Visualization:** The correlation matrix is visualized using a heatmap, where colors represent the magnitude and direction of correlations. Annotated values allow for easy identification of strong positive or negative relationships.

Comparing the correlation matrices of the original imbalanced dataset and the SMOTE + ROS balanced dataset provides additional insights into how resampling affects feature relationships and ensures that balancing does not introduce unintended biases.

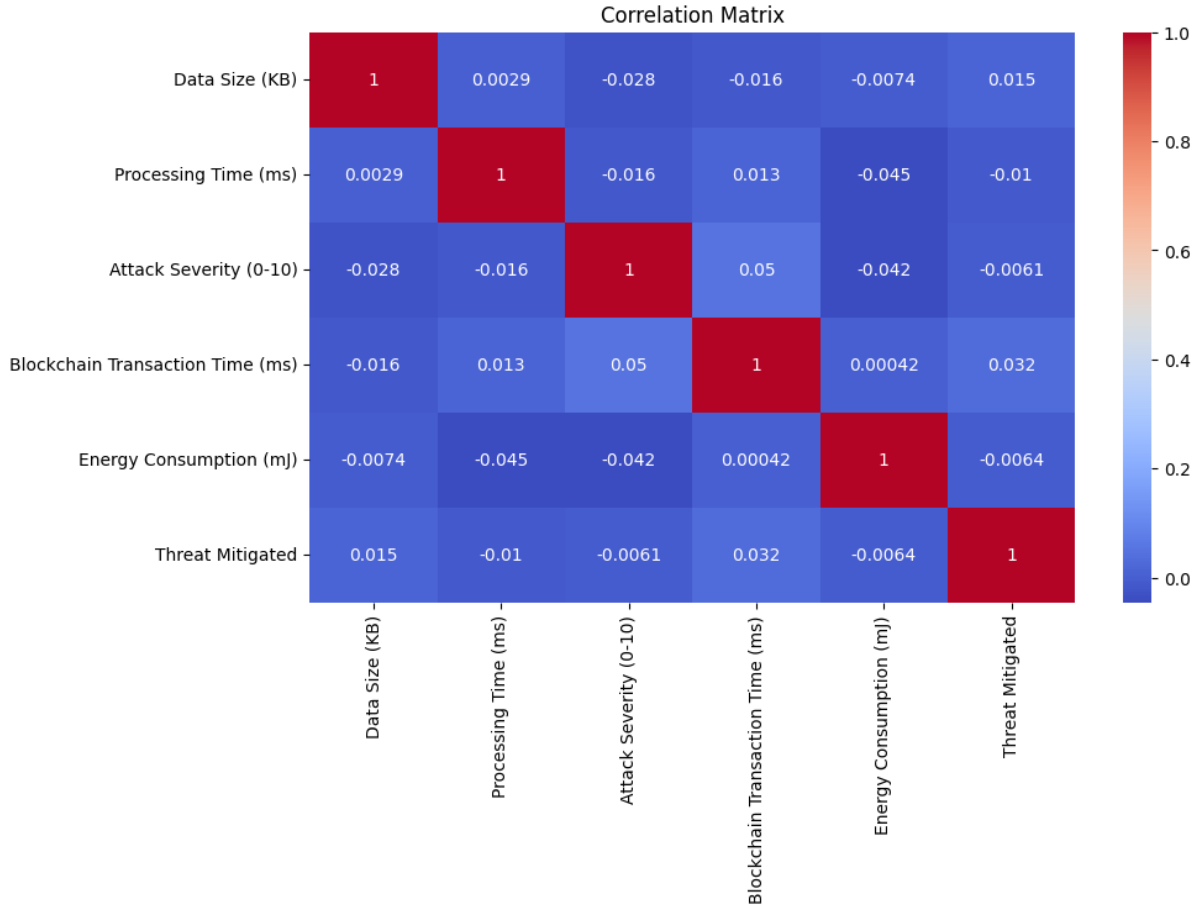


Fig 3.2.1.6 Correlation Matrix

3.2.2 Handling Class Imbalance

The dataset used in this study exhibits a significant imbalance in the target variable Threat Mitigated, where the majority class (0 – Not Mitigated) constitutes 71.7% of the samples, while the minority class (1 – Mitigated) accounts for only 28.3%. Such an imbalance poses a critical challenge in machine learning, as most algorithms assume roughly balanced class distributions.

Importance of Addressing Class Imbalance

If left unaddressed, class imbalance can severely bias the model toward the majority class, resulting in high overall accuracy but poor predictive performance for the minority class. This is especially problematic in security and IoT contexts, where the minority class (mitigated threats) is often the more important outcome. A biased model may fail to correctly identify mitigated threats, leading to ineffective security measures.

Effects on Machine Learning Models

I. **Decision Boundaries**: Imbalanced data can skew decision boundaries toward the majority class, reducing the model's sensitivity to the minority class.

II. **Evaluation Metrics Misleading**: Metrics like accuracy may appear high because the model predicts the majority class correctly most of the time, even if it fails to detect minority instances. Metrics such as precision, recall, F1-score, and ROC-AUC are more informative in such cases.

III. **Overfitting to Majority Class**: The model may learn patterns only for the majority class, ignoring subtle but important patterns in the minority class.

Techniques for Handling Class Imbalance

To mitigate these issues, various resampling techniques are employed:

I. **Random Undersampling (RUS):** This process reduces the number of majority class samples to match the minority class. This process balances the dataset, but it can lead to the loss of potentially valuable information.

II. **Random Oversampling (ROS):** Samples of the minority classes are duplicated in order to balance the dataset. ROS also retains all the majority class information but can bring overfitting due to repetition of the same minority samples.

III. **Synthetic Minority Oversampling Technique (SMOTE):** Seeks to create artificial samples of the minority group by interpolating existing minority examples. The SMOTE expands the diversity of minority classes and the model generalizes more than in the case of simple oversampling.

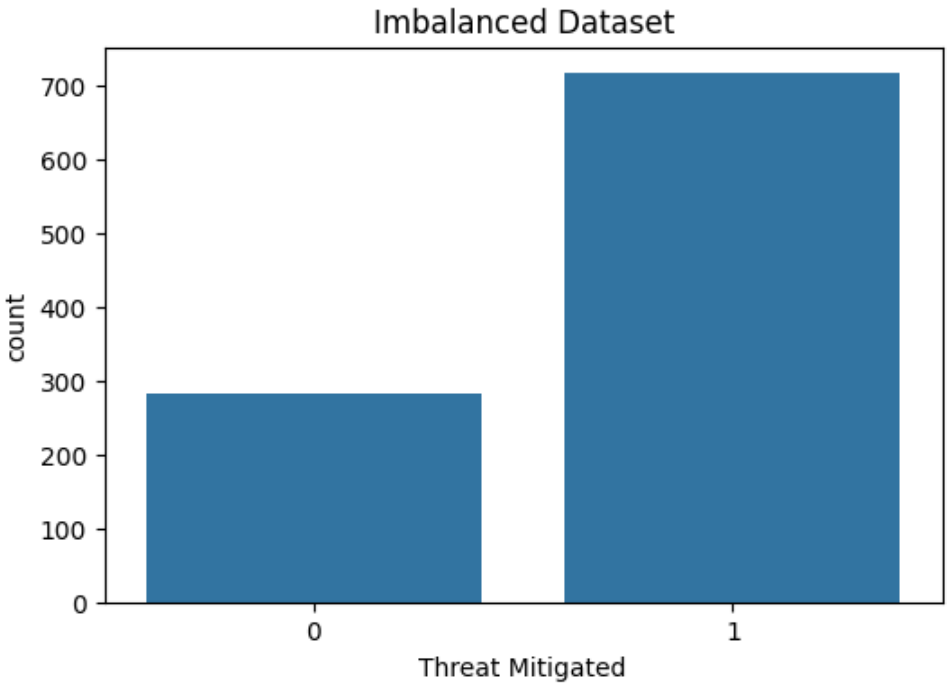


Fig 3.2.2.1 Imbalance Class Distribution of Threat Mitigated

Selected Approach

A hybrid method that comprised of SMOTE and Random Oversampling (ROS) was employed in this study. The first one is that SMOTE creates a wide-ranging synthetic minority samples, and the second is that ROS replicates more minority examples (when necessary) to have a completely balanced dataset:

- Preservation of the majority class information
- Enhanced representation of the minority class
- Robust and unbiased model training

By applying this hybrid resampling method, the dataset becomes balanced and suitable for machine learning, allowing models to accurately learn from both classes and improving performance metrics such as recall and F1-score for the minority class.

Approaches considered:

- **Undersampling the majority class:** The majority class will be under-represented to have an equivalent number of samples as the minority class. Threat: absence of valuable information.

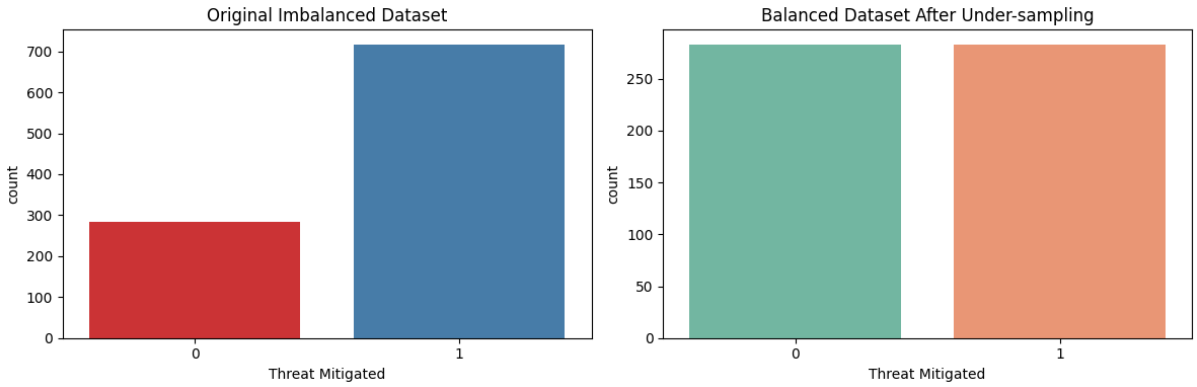


Fig 3.2.2.2 Under-sampling the majority class

- **Random Oversampling of the minority class:** Duplicates minority samples. Risk: potential overfitting.

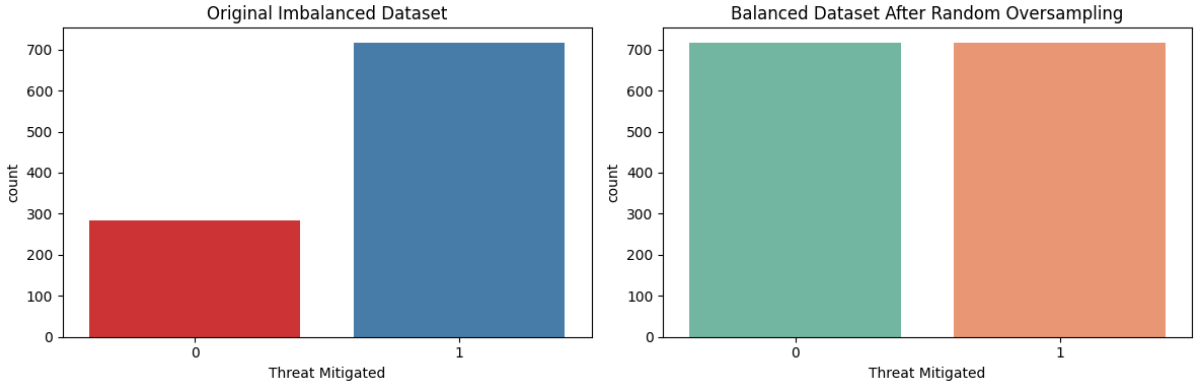


Fig 3.2.2.3 Random Oversampling of the minority class

- **Synthetic Oversampling (SMOTE):** Generates synthetic minority samples using interpolation.

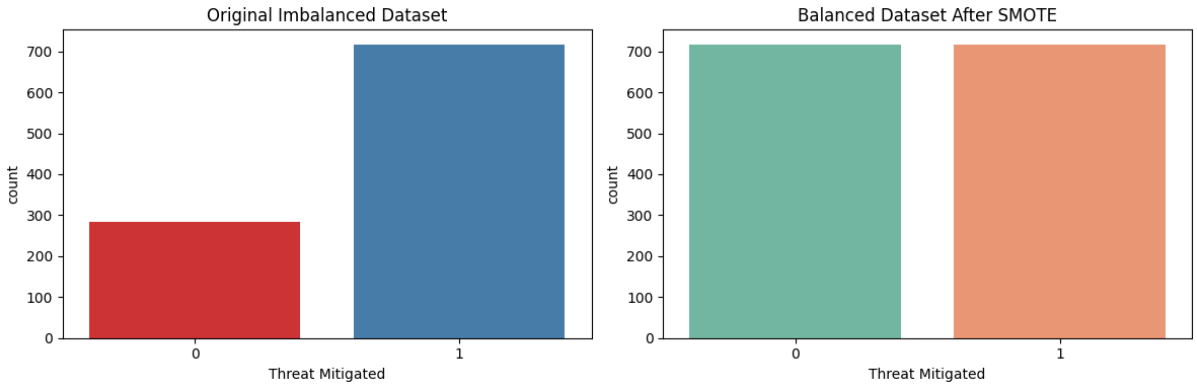


Fig 3.2.2.4 Synthetic Oversampling (SMOTE)

Selected Approach: To manage successfully the imbalance of classes in the IoT security dataset, the combination of SMOTE and Random Oversampling (ROS) was used. The SMOTE is initially used to sample the minority classes synthetically so that they are diverse and at risk of overfitting since they are already similar. Subsequently, ROS is used to further balance the classes by replicating the instances where needed, arriving at the same number of the all target classes. This hybrid approach provides the benefits of both approaches, namely synthetic sample generation and controlled oversampling, to produce a well-formed and balanced dataset to be used in machine learning model training and evaluation.

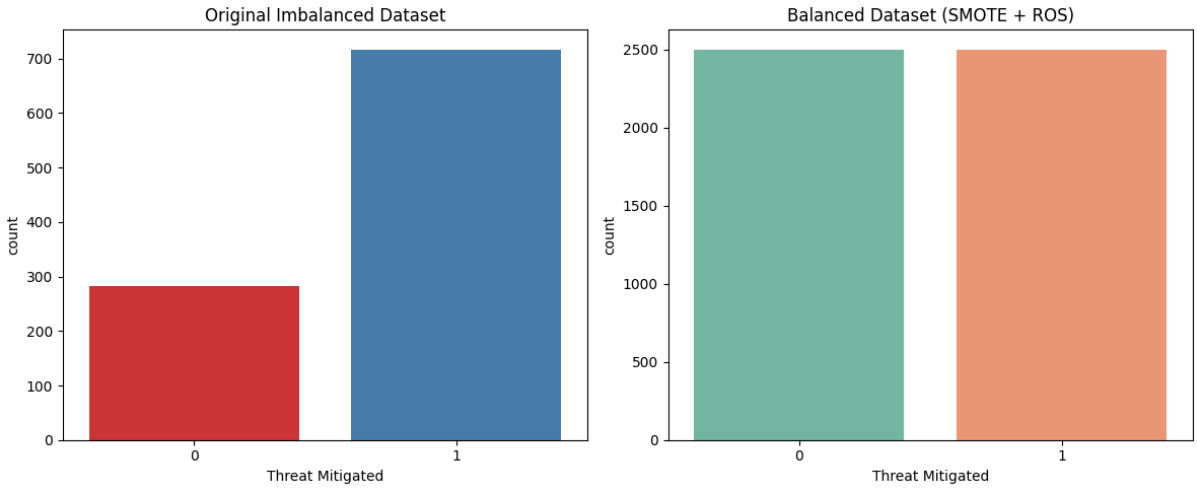


Fig 3.2.2.5 Combining SMOTE and Random Oversampling (ROS)

SMOTE Formula:

$$X_{new} = X_i + \lambda * (X_{zi} - X_i), \qquad \lambda \in [0,1]$$

- This approach generates a new sample along the line connecting Xi and Xzi.
- After SMOTE generates synthetic samples then ROS duplicates the majority class instances to reach equal class sizes.

Justification:

The avoidance of undersampling was due to the fact that it would delete information of value to the majority class thus, weakening the performance of the model. Repeated instances can occur solely through random oversampling and this raises the chances of overfitting. With the help of SMOTE, artificial samples of the minority category are created, which adds diversity and maintains the original distribution of data. The SMOTE model in combination with the Random Oversampling (ROS) can provide a completely balanced dataset and can offer a significant basis in which models can be trained and predictive results will be more accurate.

Summary of Imbalanced vs Balanced Dataset

Dataset	Class 0 Count	Class 1 Count	Total Samples	Class 0 %	Class 1 %
Imbalanced	283	717	1000	28.3	71.7
Balanced	2500	2500	5000	50.0	50.0

Fig 3.2.2.6 Summary of Balance and Imbalance Dataset

3.2.3 Encoding Categorical Variables

Categorical variables in the dataset need to be transformed into numerical formats for machine learning algorithms. Two main approaches were considered:

- **Label Encoding:** Converts each category into a unique integer. For a categorical feature X with categories {C1, C2, ..., Cn, Label Encoding maps each category Ci → i.
- **One-Hot Encoding (OHE):** Represents each category as a binary vector. For a feature X with n categories, OHE creates n−1 new binary columns (to avoid multicollinearity), where each column indicates the presence (1) or absence (0) of a category.

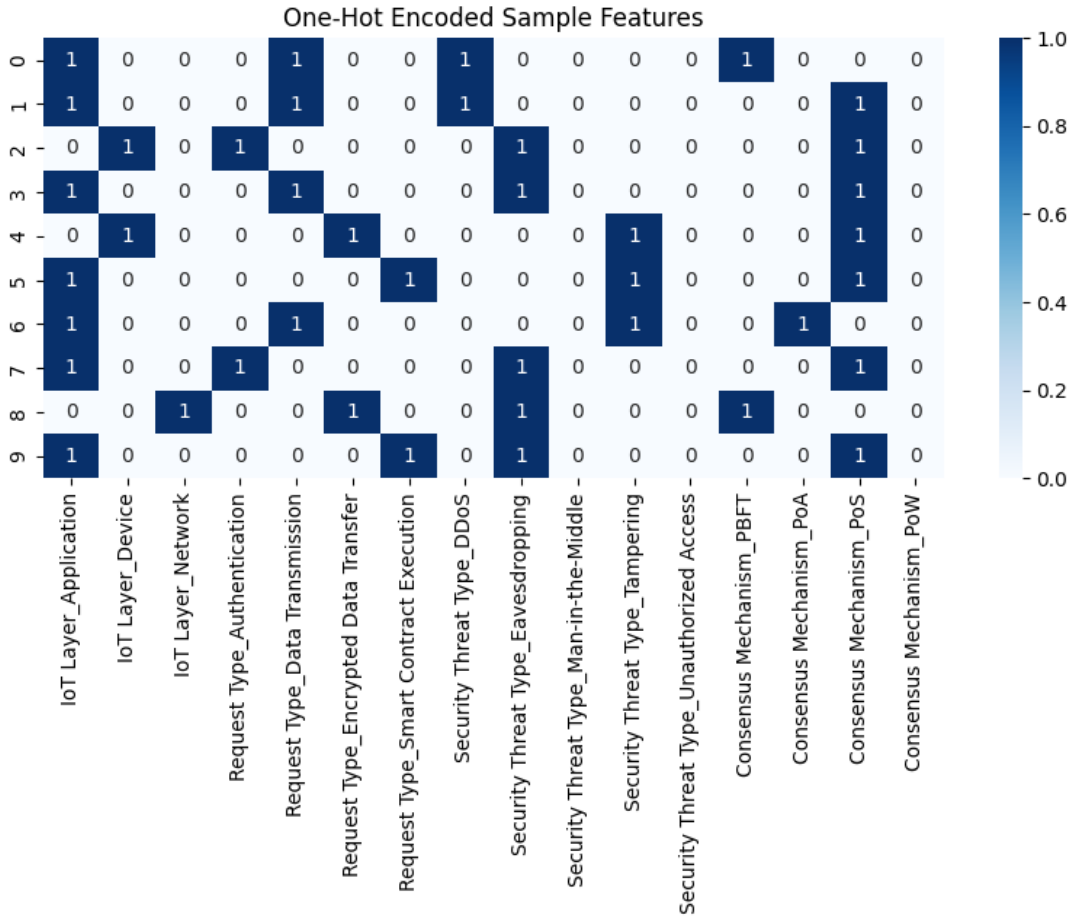


Fig 3.2.3.1 One-Hot Encoding

Formula (OHE representation):

$$OHE(Xi) = \begin{cases} 1, & if\ X = Ci \\ 0, & otherwise \end{cases}$$

Selected Approach

To prepare the dataset for machine learning models, appropriate encoding techniques were applied to transform categorical features into numerical representations. The steps are as follows:

Label Encoding:

- Applied to identify the binary target variable (Threat Mitigated).
- Label encoding converts each unique category into a numeric label.

For the target variable, this transformation allows the machine learning algorithms to interpret the output as a numeric class for classification purposes.

One-Hot Encoding (OHE):

- Applied to nominal categorical features such as IoT Layer, Request Type, Security Threat Type, and Consensus Mechanism.
- One-Hot Encoding creates new binary columns for each category of a feature (e.g., IoT Layer: Application → [1, 0, 0], Device → [0, 1, 0]).

This approach ensures that the model does not assume any ordinal relationship among categories, preserving the nominal nature of the features.

Justification

Label Encoding:

Label encoding is computationally effective and applicable in unique identifiers or target variables that have two classes only. Nevertheless, it can also create that illusion of sanity when used on nominal features with more than two categories (e.g. IoT Layer), which can lead the model to make assumptions that the numbers can be read as the number of levels.

One-Hot Encoding:

One-Hot Encoding averts the occurrence of this ordinal misinterpretation by describing each category to individual binary columns. This will make sure that there is no ranking or precedence of any categories. Therefore, the model will be able to learn significant patterns only when a category occurs or not..

Label Encoding identifiers and binary outputs combined with One-Hot Encoding nominal features numerically represent the dataset in a manner that is both computationally efficient and compatible with machine learning algorithms and limits the possibility of bias or incorrect assumptions regarding the relationships between features.

3.2.4 Scaling Numerical Features

Machine learning algorithms often require features to be on a similar scale to prevent features with large numerical ranges from dominating the model training. In this study, the numerical features such as Data Size, Processing Time, Attack Severity, Blockchain Transaction Time, and Energy Consumption were scaled using Min-Max normalization and Z-score standardization. Two scaling methods were considered:

Min–Max Normalization:

$$X' = \frac{(X - X_{min})}{X_{max} - X_{min}}$$

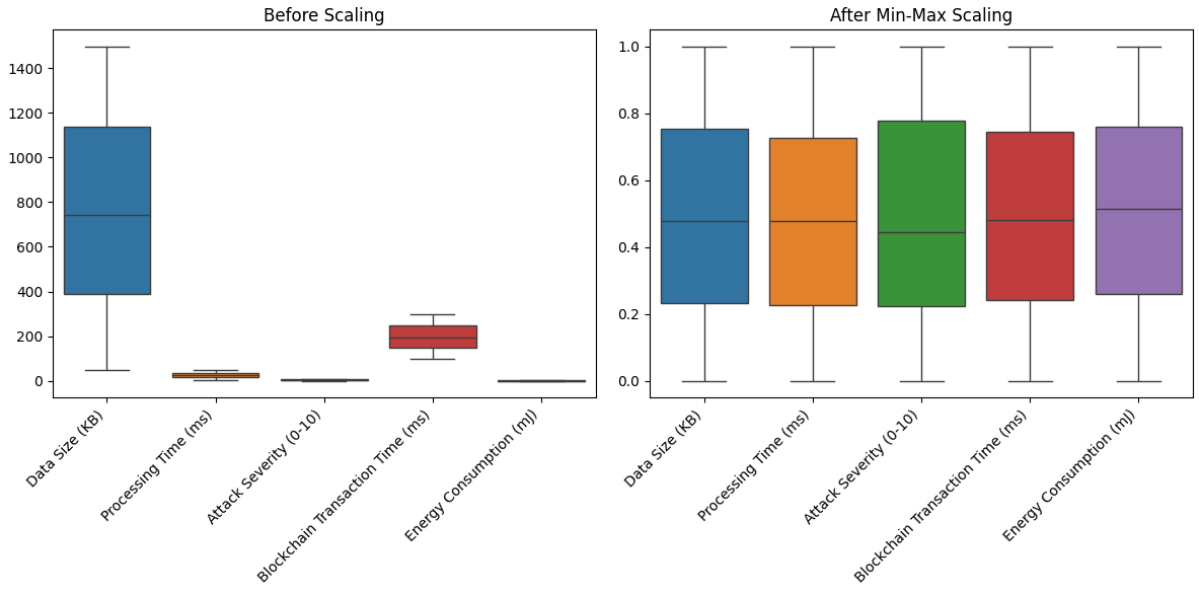


Fig 3.2.4.1 Min-Max Normalization

Where:

- X is the original value,
- X_{min} and X_{max} are the minimum and maximum values of the feature,
- X' is the scaled value.

Purpose and Advantages:

- Transforms all numerical features to the same range, ensuring that no single feature dominates due to scale differences.
- Useful for algorithms that rely on distance measures, such as k-Nearest Neighbors (k-NN) and Support Vector Machines (SVMs).

Z-score Standardization (selected):

Z-score standardization (also called standard scaling) transforms numerical features so that they have a mean of 0 and a standard deviation of 1. The formula is:

$$X' = \frac{(X - \mu)}{\sigma}$$

Where:

- X = original feature value
- μ= mean of the feature
- σ = standard deviation of the feature
- X′ = standardized value

Purpose of Z-Score Standardization

The primary purpose of Z-score standardization is to transform numerical features to a common scale without distorting differences in the ranges of values. This ensures that:

1. All features contribute equally during model training.
2. Algorithms that rely on distance measures or gradient-based optimization function properly.
3. Features with larger magnitudes do not dominate the learning process, which could bias the model.

Advantages

1. **Centers data at zero:**
    - By subtracting the mean (X− μ), the transformed feature has a mean of 0, which improves model convergence.
2. **Scales to unit variance:**
    - Dividing by standard deviation (σ) ensures all features have a variance of 1, making them comparable.
3. **Handles outliers better than Min-Max normalization:**
    - Extreme values do not compress the majority of the data into a small range.

4. **Improves performance of scale-sensitive algorithms:**
    - Algorithms like Support Vector Machines (SVMs), K-Nearest Neighbors (KNN), logistic regression, and neural networks benefit from standardized features.
5. **Facilitates faster convergence:**
    - Standardized features help optimize algorithms (e.g., gradient descent) to converge faster and more reliably.

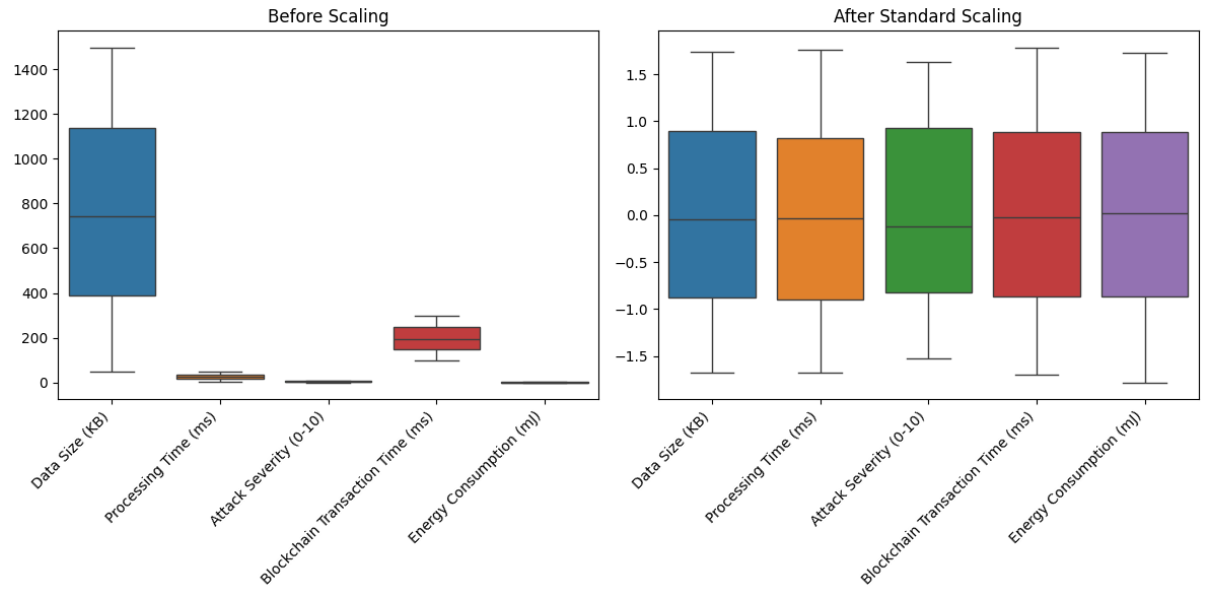


Fig 3.2.4.2 Z-score standardization

Justification:

Z-score standardization was chosen over Min-Max normalization because it maintains the relative distribution of the data, centers each feature around zero mean, and scales the variance to one. Unlike Min-Max scaling, it does not compress the feature values due to outliers, ensuring that meaningful variations in the data are preserved. This makes it more suitable for algorithms that are sensitive to feature scaling and distribution assumptions, such as Principal Component Analysis (PCA), Support Vector Machines (SVM), and Neural Networks, where balanced feature contributions are essential for effective learning.

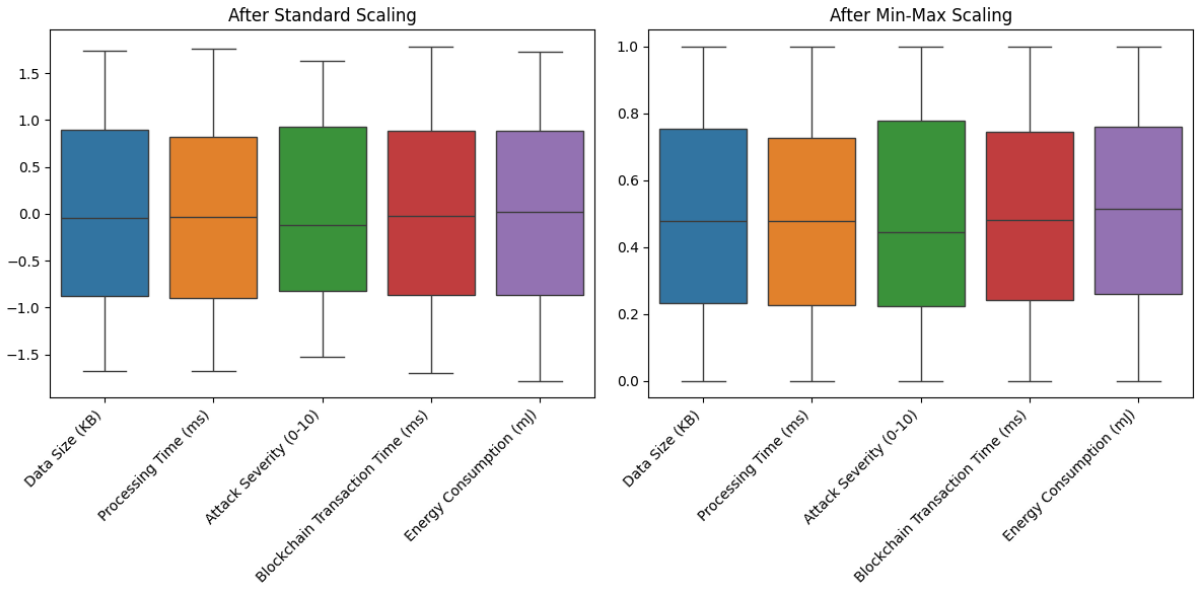
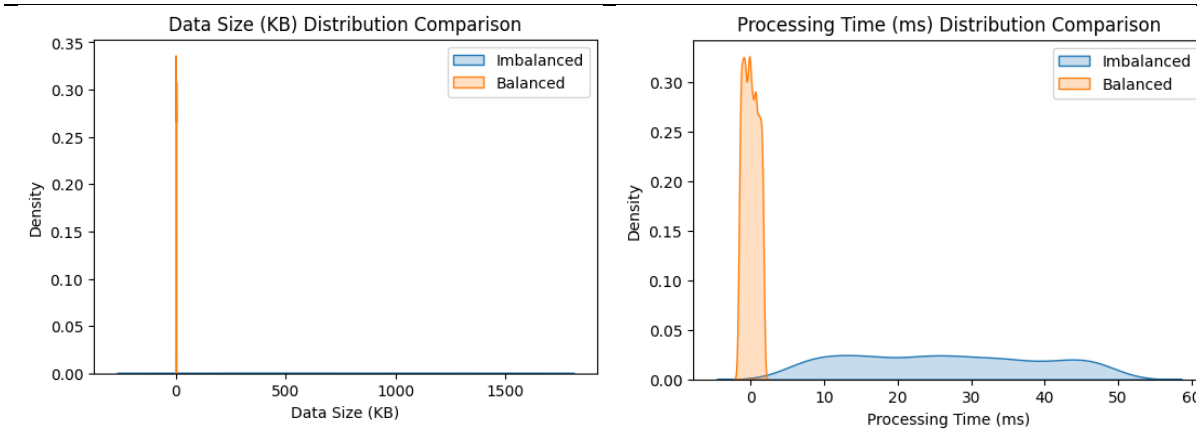


Fig 3.2.4.3 Standard Scaling and Min-Max scaling highlights

3.2.5 Class Distribution of Balanced Dataset After Preprocessing

After applying preprocessing and class balancing techniques, including SMOTE and Random Oversampling (ROS), the dataset became perfectly balanced with respect to the target variable Threat Mitigated. Both classes, labeled 0 and 1, now contain an equal number of samples, ensuring that the machine learning models trained on this dataset will not be biased toward any class. This balanced distribution provides a solid foundation for robust and reliable predictive analysis.



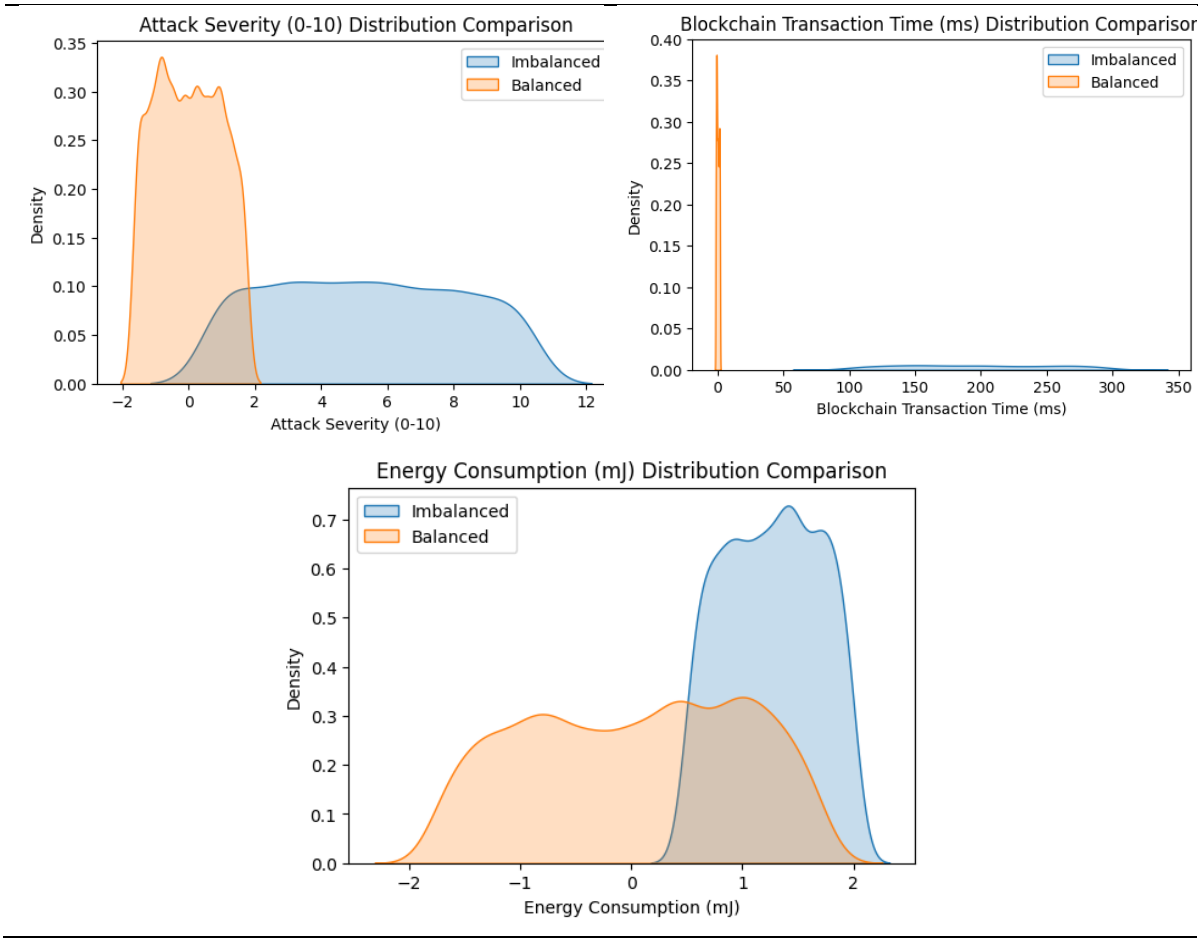


Fig 3.2.5.1 Class Distribution of Categorical Data after Preprocessing Data

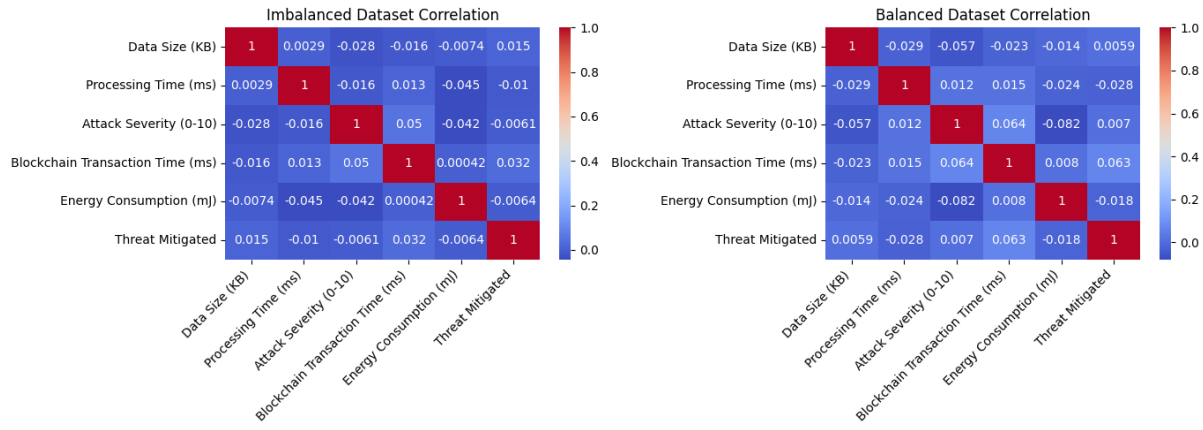


Fig 3.2.5.2 Correlation Matrix of Imbalance and Balance Data

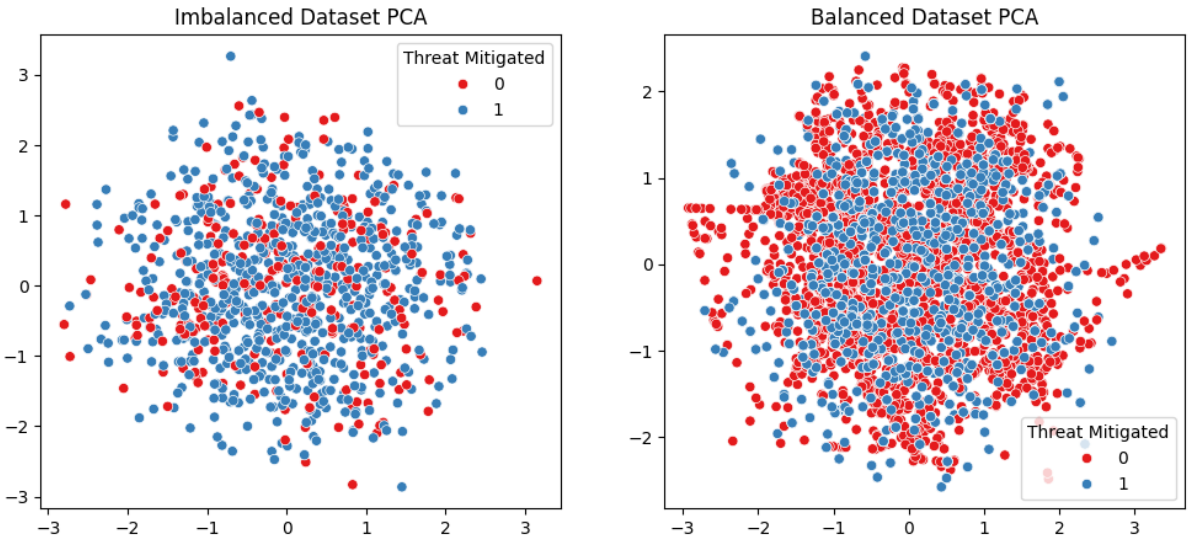


Fig 3.2.5.3 PCA Matrix of Imbalance and Balance Data

In this study, PCA was applied to the balanced IoT dataset to visualize the separation of classes (Threat Mitigated = 0 or 1) in a 2D scatter plot. This helps in understanding the effect of preprocessing and balancing on the dataset.

3.2.6 Dataset Reconstruction and Export

After preprocessing and resampling:

- All numerical and encoded categorical features were recombined into a single structured dataset.
- The target variable (*Threat Mitigated*) was appended.
- The dataset was exported in CSV format for reproducibility and downstream model development.

3.3 Project Plan

The project plan outlines the complete workflow of the study, starting with dataset collection, where raw IoT security data is gathered, examined for quality, and prepared for further processing. In the data preprocessing stage, identifiers such as Device ID are label-encoded, categorical features are one-hot encoded, and numerical features are scaled to standardize the dataset. This ensures that the data is clean, consistent, and suitable for machine learning models.

Following preprocessing, class imbalance handling is performed using SMOTE and Random Oversampling (ROS) to create a balanced dataset. This step is crucial to prevent bias toward the majority class and to improve the performance of classifiers. The balanced data is then used in model training, where various machine learning algorithms—including Logistic Regression, Decision Tree, Random Forest, Gradient Boosting,

XGBoost, LightGBM, SVM, KNN, Naive Bayes, and Neural Networks—are applied to predict the target variable.

After training, model evaluation is conducted using metrics such as accuracy, precision, recall, and F1-score to measure performance. The workflow concludes with a comparative analysis to select the best-performing model. This structured approach provides clarity on each stage of the study, highlights decision points, and visually represents the methodology for reproducibility and understanding.

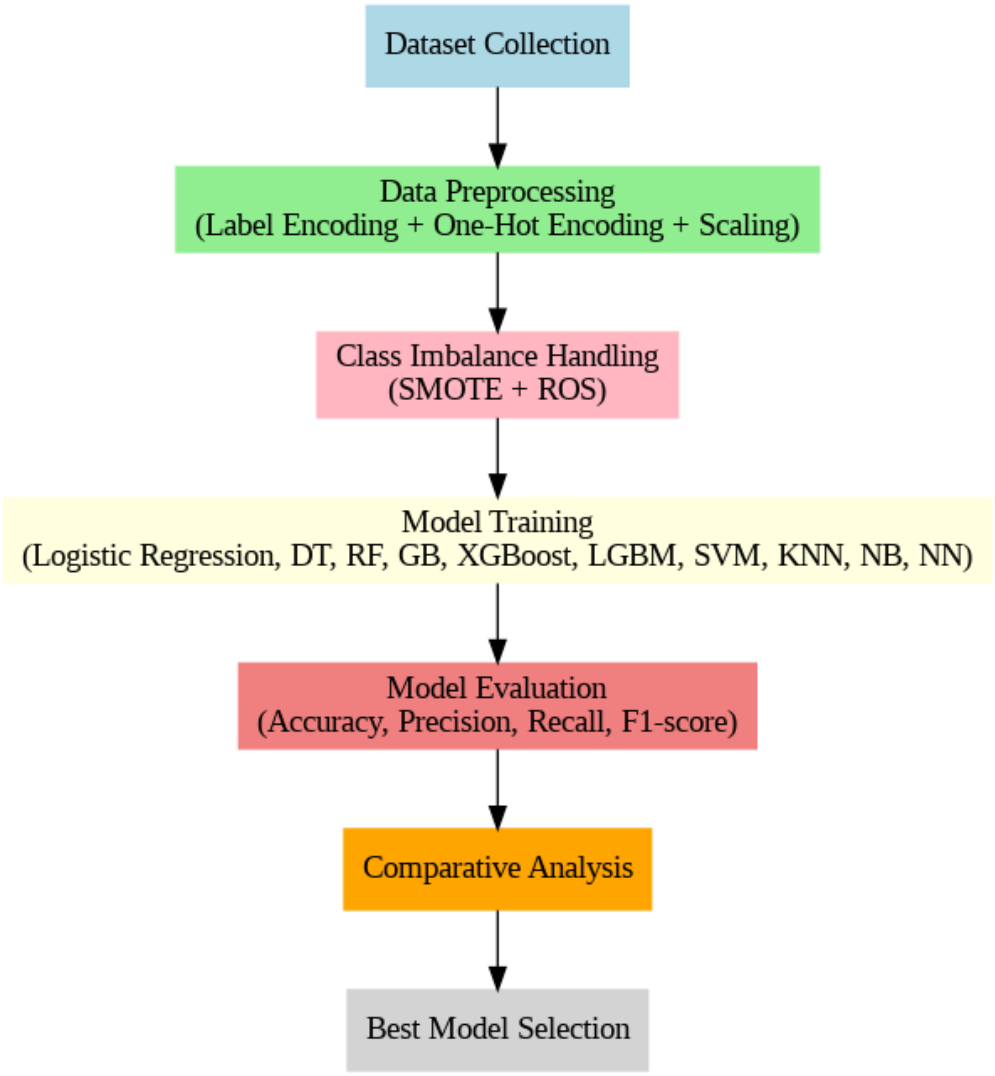


Fig 3.3.1 Overall Project Plan

3.4 Task Allocation

This table depicts the timeline of the principal activities in each period of the project. from week 12 to week 48

Tasks	Weeks																		
	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48
Data Collection Phase																			
Preprocess all the data																			
Model training																			
Evaluate, Finalize Models.																			

3.5 Summary

The methodology of this study is designed to systematically preprocess, balance, and analyze an IoT security dataset to facilitate accurate threat detection using machine learning. The process begins with requirement analysis, where the objectives of the study are defined, the relevant features and target variable ("Threat Mitigated") are identified, and the dataset is assessed for completeness and consistency. This ensures that the data is suitable for subsequent preprocessing and modeling tasks.

In the data collection and preparation stage, the dataset is imported from the source, and identifiers such as Device ID are label-encoded. Nominal categorical features—including IoT Layer, Request Type, Security Threat Type, and Consensus Mechanism—are one-hot encoded to preserve their categorical nature. Numerical features, such as Data Size, Processing Time, Attack Severity, Blockchain Transaction Time, and Energy Consumption, are standardized using Z-score normalization, which centers the data at zero mean and scales the variance to one. This step ensures that all features contribute equally to the learning process and prevents any single feature from dominating the model due to scale differences.

The dataset exhibits a significant class imbalance in the target variable, which could bias machine learning models toward the majority class. To address this, a hybrid approach combining SMOTE (Synthetic Minority Oversampling Technique) and Random Oversampling (ROS) is applied. By creating artificial samples that belong to the minority group, SMOTE makes it more varied and prevents the exact duplicates, whereas the ROS attempts to even out the remaining disparities by replicating the samples of the minority group and creating a balance. This makes sure that the dataset is non-bias and fits well in strong model training.

After the preprocessing and balancing, there will be the methodology, whereby, the processed features are joined together format into a structured DataFrame with the features clearly named, then the target variable is added. The analysis of the relationships between the features, the interpretation of the impact of the preprocessing, and the validation of the balancing process are analyzed using visualization methods, i.e. class distribution plots, correlation matrices, histograms, scatterplot matrices (pair plots), PCA projections.

Lastly, the model training and evaluation are also part of the methodology, and the various machine learning algorithms, such as Logistic Regression, Decision Tree, Random Forest, Gradient Boosting, XGBoost, LightGBM, SVM, KNN, Neural Networks, and Naive Bayes, are also trained and evaluated. Accuracy, precision, recall and F1-score performance measures are derived and the results compared to determine the most performing model. Such systematic and elaborated design will guarantee reproducibility, give good understanding of the preprocessing and modeling pipeline and establish a robust basis of threat detection in the IoT environments.

Chapter 4

Implementation and Results

The chapter outlines the implementation procedure, evaluation plans, and performance results of the suggested methodology. It explains the setting of the environment to be used in the experiment, the methods to be used in testing and evaluation of performance measurement and provides the discussion over the results obtained and comparative insights.

4.1 Environment Setup

This study used the Google Colab as a cloud computing platform to perform its experiments, which is a powerful and interactive environment to execute Python code without having to use local computing resources. The choice of Google Colab was facilitated by its simplicity, the free accessibility to the use of the GPU/TPU acceleration, and its integration with Google Drive so that the datasets could be loaded and stored directly. Reproducibility, teamwork and efficient experimentation, which is essential to machine learning projects, was made easier by the environment.

The Python 3.11 programming language was used that is compatible with a vast array of recent libraries to analyze, perform machine learning, and visualize data. The basic libraries were Pandas and NumPy to do data manipulation, cleaning, and numerical computations, Scikit-learn to do preprocessing, model construction and evaluation, and Imbalanced-learn (imblearn) to deal with class imbalance with methods such as SMOTE and Random Oversampling. Data distribution, correlation, and preprocessing visualization were done using Matplotlib and Seaborn which offered the quality of plotting data to explore and report data.

To apply more sophisticated machine learning models, XGBoost and LightGBM libraries were used as the gradient boosting algorithms, and the TensorFlow/Keras was used to design and model a deep neural network. Several different algorithms including traditional classifiers such as Logistic Regression and Decision Trees, as well as ensemble and neural network methods, could be evaluated using these tools, which allowed conducting a comparative analysis across all of them.

The data was uploaded to Google Drive and then directly loaded into Colab to make the integration as seamless as possible and remove the problem of local storage capacity. Preprocessing pipeline consisted of identifier feature labelling, nominal categorical feature one-hot encoding, numerical feature scaling with Z-score standardization, SMOTE and Random Oversampling to balance the proportions of the classes. Through Colab, these preprocessing steps and the further modeling workflows were performed efficiently, and controlled random seeds provided reproducible results in a variety of experiments.

In general, the Colab platform offered a scalable, powerful, and easy to use environment which allowed conducting systematic experiments, effective computations and clear visualization of the raw and processed IoT security dataset, which was fundamental to assessing the efficacy of the proposed methodology.

4.2 Testing and Evaluation

In this work, following the preprocessing and balancing of the IoT security dataset, several model machine learning models were trained and tested. Testing and evaluation aims to evaluate the predictive power of these models and find the best approach that predicts the most effective way of identifying whether a threat is mitigated or not. Both classic machine learning algorithm (Logistic Regression, Decision Trees, Random Forests, Gradient Boosting, SVM, KNN, Naive bayes) and more progressive ensemble and deep learning algorithms (XGBoost, LightGBM, Neural Networks) were taken into account.

4.2.1 Evaluation Metrics

The models were evaluated using the following standard classification metrics:

1. Accuracy

Accuracy measures the proportion of correct predictions among the total predictions. It is defined as:

$$Accuracy = \frac{(TP + TN)}{TP + TN + FP + FN}$$

where:
- TP = True Positives
- TN = True Negatives
- FP = False Positives
- FN = False Negatives

2. Precision

Precision measures the proportion of correctly predicted positive observations among all predicted positive

$$Precision = \frac{TP}{TP + FP}$$

Recall (Sensitivity)

Recall indicates the proportion of correctly predicted positive observations among all actual positives:

$$Recall = \frac{TP}{TP + FN}$$

F1-Score

The F1-score is the harmonic mean of precision and recall, providing a balance between the two metrics:

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

These metrics were chosen because the dataset exhibits initial class imbalance, making accuracy alone insufficient. Precision, recall, and F1-score provide a more robust evaluation of model performance, particularly for minority classes.

Specificity

$$Specificity = \frac{TN}{TN + FP}$$

4.2.2 Comparative Analysis of Models

The evaluation was performed on both the imbalanced dataset and the balanced dataset after applying SMOTE and Random Oversampling. Comparative analysis revealed:

Logistic Regression

Logistic Regression is a linear classifier that models the probability of the target variable using the logistic function:

Equation:

$$P(Y = 1|X) = \frac{1}{1 + e^{-(\beta 0 + \beta 1 * X1 + \beta 2 * X2 + \ldots + \beta n * Xn)}}$$

It works well for linearly separable datasets, but is limited in capturing complex nonlinear relationships. On imbalanced datasets, it tends to be biased toward the majority class.

Decision Tree: Decision Trees are non-linear models that recursively split the feature space based on thresholds that maximize information gain or Gini impurity.

Equation for Gini Impurity:

$$Gini = 1 - \Sigma (p_i)^2$$

They are intuitive and easy to interpret but prone to overfitting, especially on small or noisy datasets. Proper pruning is necessary to maintain generalization.

Random Forest

Random Forest is an ensemble method combining multiple Decision Trees using bagging and feature randomness to improve generalization. It reduces overfitting and increases stability compared to a single Decision Tree. The final prediction is based on majority voting for classification:

Equation:

$$y\_pred = mode\left(y_{pred}^{1}, y_{pred}^{2}, \ldots, y_{pred}^{n}\right)$$

Gradient Boosting, XGBoost, LightGBM

These are boosting algorithms that sequentially build weak learners to correct the errors of previous models. They are highly effective for complex datasets, capturing intricate relationships between features.

Boosting update rule:

$$\boldsymbol{F_{m(x)} = F_{\{m-1\}(x)} + \eta \cdot h_{m(x)}}$$

Where $F_{m(x)}$ is the ensemble prediction at iteration m, $h_{m(x)}$ is the new weak learner, and η is the learning rate.

Support Vector Machine (SVM)

SVM seeks an optimal hyperplane that maximizes the margin between classes. It is effective for high-dimensional data and can handle non-linear relationships using kernel functions:

Decision function:

$$f(x) = sign(\Sigma\, \alpha_i * y_i * K(x_i, x) + b)$$

Where K (x_i , x) is the kernel function, α_i are the Lagrange multipliers, and b is the bias term.

K-Nearest Neighbors (KNN)

KNN classifies a sample based on the majority label among its k-nearest neighbors. It is simple and non-parametric, but sensitive to the choice of k and feature scaling. Distance metrics like Euclidean distance are typically used:

Euclidean Distance:

$$d(x, y) = sqrt(\Sigma (x_i - y_i)^2)$$

Naive Bayes

Naive Bayes assumes conditional independence between features and applies Bayes' theorem for classification. Its performance is limited when features are correlated.

Bayes' theorem:

$$P(C|X) = P(X|C) * \frac{P(C)}{P(X)}$$

Neural Network

Neural Networks consist of multiple layers of interconnected neurons capable of learning complex nonlinear relationships. They are flexible and highly effective on balanced datasets. The output of a neuron is calculated as:

Neuron output:

$$y = f(\Sigma w_i * x_i + b)$$

Where **f** is the activation function, $\boldsymbol{w_i}$ are weights, $\boldsymbol{x_i}$ are inputs, and b is the bias.

The comparative analysis demonstrates that tree-based ensembles and neural networks are particularly effective for the IoT threat prediction problem, especially when the dataset is balanced. Linear models like Logistic Regression are limited by their inability to capture non-linear patterns, while distance-based and probabilistic models are sensitive to feature scaling and assumptions. Choosing an appropriate algorithm depends on data complexity, feature relationships, and the need for interpretability versus predictive power.

The balanced dataset consistently improved recall and F1-score for the minority class compared to the imbalanced dataset. This confirms the importance of class balancing using SMOTE and ROS before model training.

4.2.3 Observations and Insights

- Ensemble models and neural networks consistently outperformed simpler models, demonstrating the importance of leveraging multiple trees or layers for complex IoT security patterns.
- Class balancing was critical to achieving high performance for the minority class, as the original imbalance could bias models toward predicting the majority class.
- F1-score was chosen as a crucial measure since it takes into consideration both the precision and recall which would provide a clearer insight into the performance of the model in terms of detecting security threats.

4.3 Results and Discussion

These experimental findings reveal the effects of the various machine learning models when used on the imbalanced and balanced data sets. The models that were trained on the imbalanced data were likely to give preference to the majority class leading to biased predictions and lack of generalization. The models became more stable and reliable when they were applied to both classes after data balancing methods (SMOTE and Random Oversampling) were used. The ensemble-based models such as the Random Forest, Gradient Boosting, XGBoost and LightGBM demonstrated good capacity in the ability to capture the interaction of multiple features, with neural networks showing superiority over other models as they could learn the non-linear decision boundaries. The simpler models like Logistic regression and Naive bayes were not able to hold the underlying patterning because they were linear and assumed that they were independent. The results indicate that proper preprocessing, data balancing, and model selection are crucial to obtaining strong threat mitigation predictions in a system of IoT security.

4.3.1 Decision Tree

Table 4.3.1.1 Classification Report of Decision Tree

Class	Precision	Recall	F1-Score	Support	Accuracy	Macro Avg	Weighted Avg
0	0.96	0.92	0.94	494	0.942	0.94	0.94
1	0.92	0.96	0.94	506	0.942	0.94	0.94

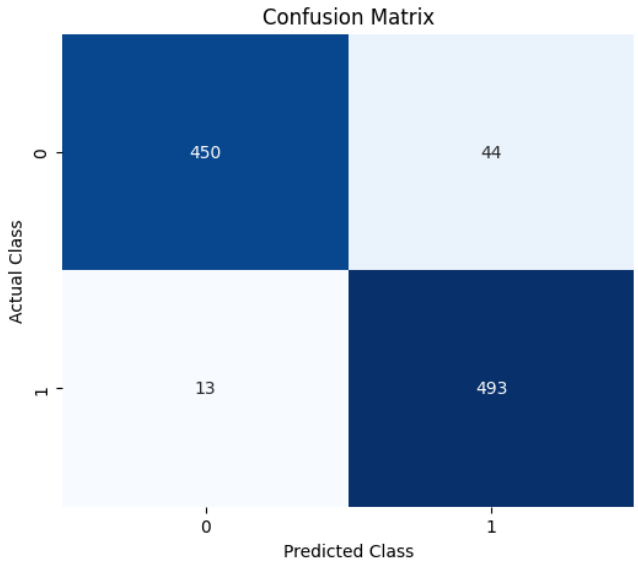


Fig 4.3.1.1 Confusion Matrix of Decision Tree

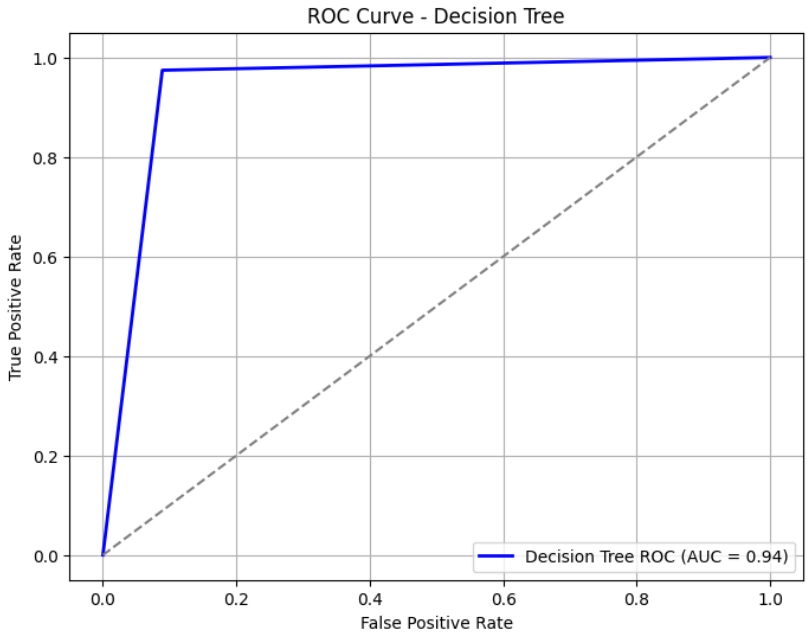


Fig 4.3.1.2 ROC Curve of Decision Tree

4.3.2 Random Forest

Table 4.3.2.1 Classification Report of Random Forest

Class	Precision	Recall	F1-Score	Support	Accuracy	Macro Avg	Weighted Avg
0	0.99	0.93	0.96	494	0.961	0.96	0.96
1	0.93	0.99	0.96	506	0.961	0.96	0.96

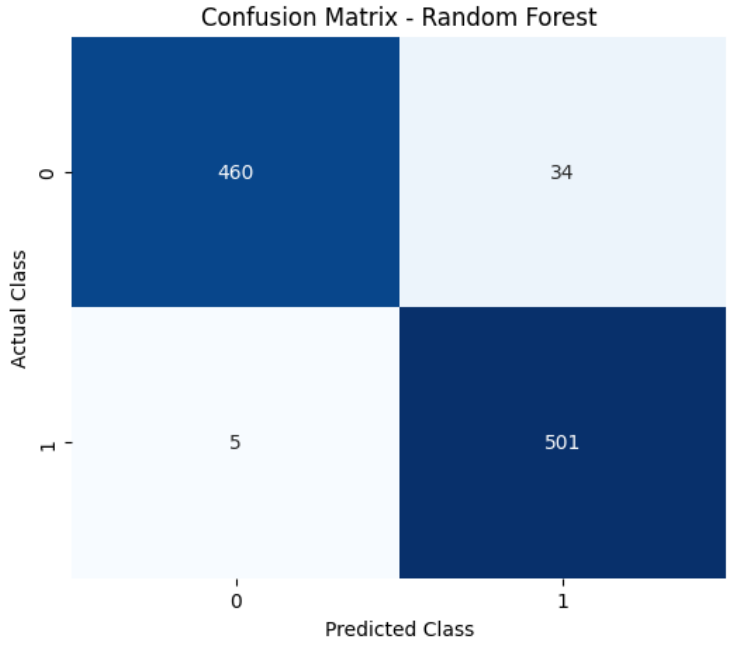


Fig 4.3.2.1 Confusion Matrix of Random Forest

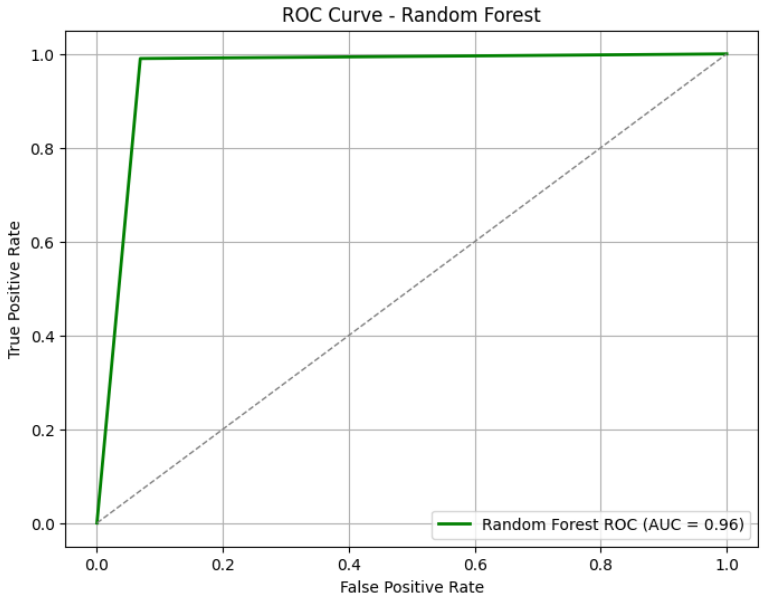


Fig 4.3.2.2 ROC Curve of Random Forest

4.3.3 Gradient Boosting

Table 4.3.3.1 Classification Report of Gradient Boosting

Class	Precision	Recall	F1-Score	Support	Accuracy	Macro Avg	Weighted Avg
0	1.00	0.83	0.91	494	0.916	0.92	0.92
1	0.86	1.00	0.92	506	0.916	0.92	0.92

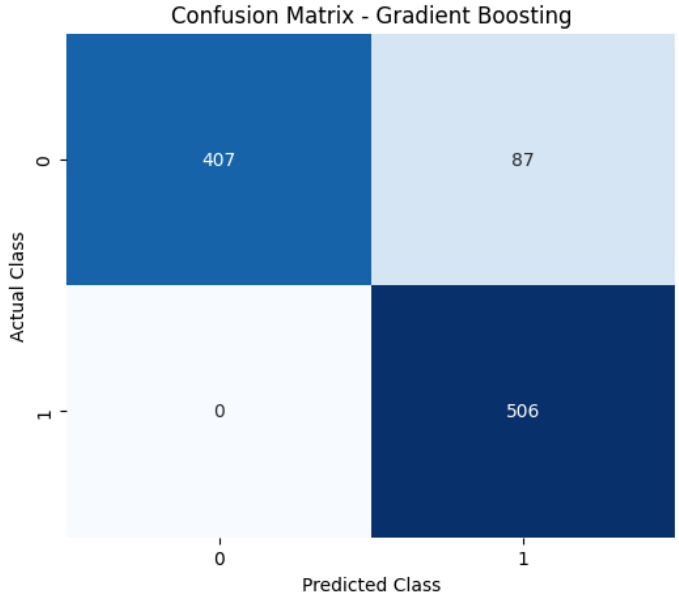


Fig 4.3.3.1 Confusion Matrix of Gradient Boosting

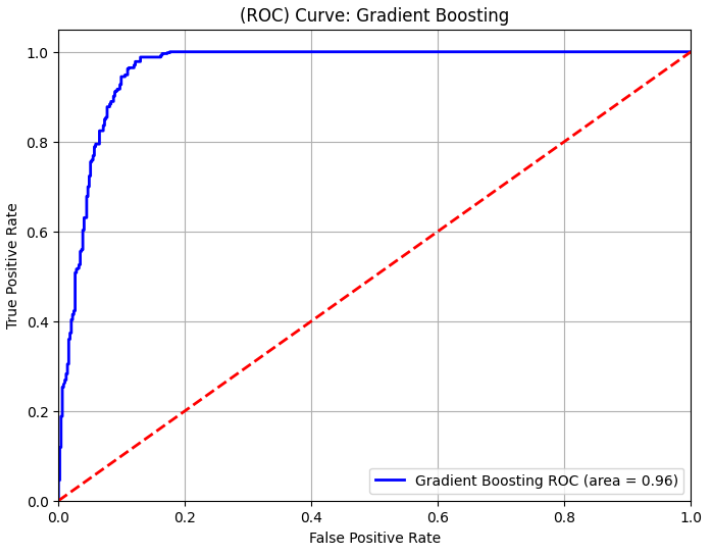


Fig 4.3.3.2 ROC Curve of Gradient Boosting

4.3.4 XG Boost

Table 4.3.4.1 Classification report of XG Boost

Class	Precision	Recall	F1-Score	Support	Accuracy	Macro Avg	Weighted Avg
0	0.99	0.90	0.94	494	0.947	0.95	0.95
1	0.91	0.99	0.95	506	0.947	0.95	0.95

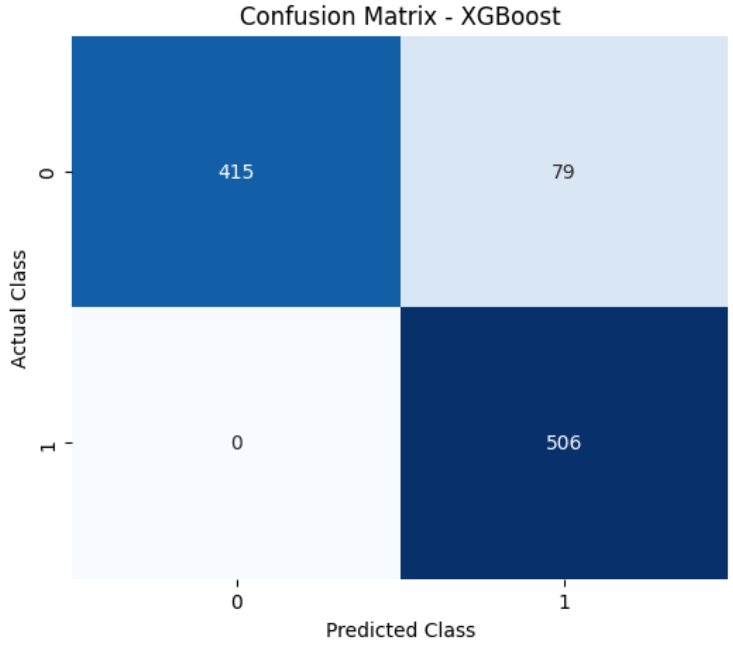


Fig 4.3.4.1 Confusion Matrix of XGBoost

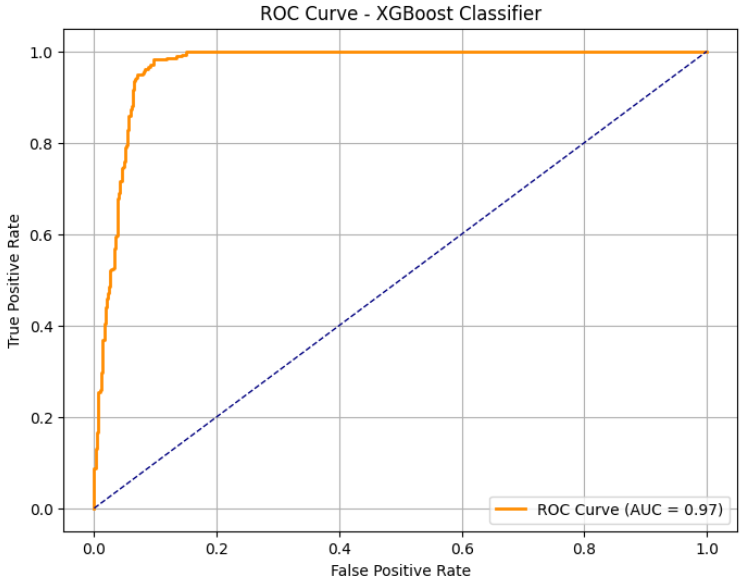


Fig 4.3.4.2 ROC Curve of XGBoost

4.3.5 Light GBM

Table 4.4.5.1 Classification Report of Light GBM

Class	Precision	Recall	F1-Score	Support	Accuracy	Macro Avg	Weighted Avg
0	0.99	0.91	0.95	494	0.95	0.95	0.95
1	0.92	0.99	0.95	506	0.95	0.95	0.95

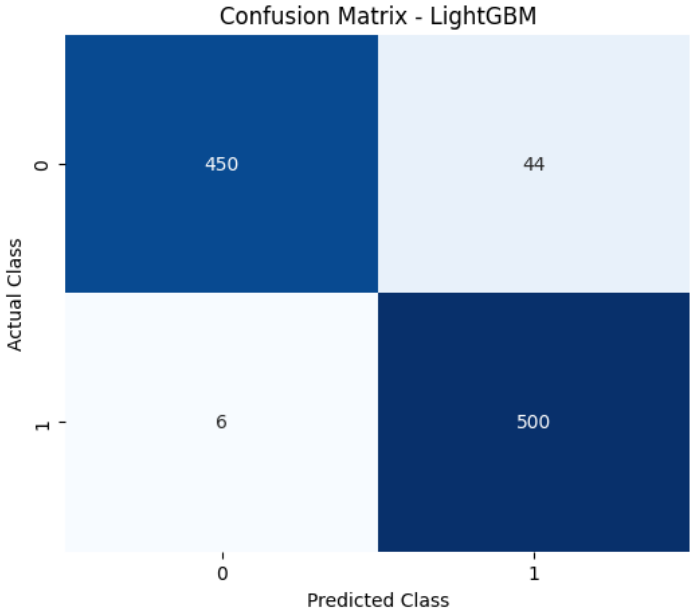


Fig 4.3.5.1 Confusion Matrix of LightGBM

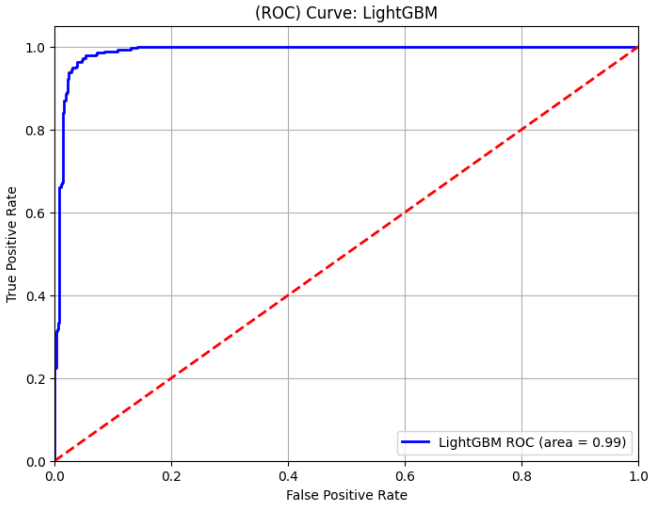


Fig 4.3.5.2 ROC Curve of LightGBM

4.3.6 SVM

Table 4.4.6.1 Classification Report of SVM

Class	Precision	Recall	F1-Score	Support	Accuracy	Macro Avg	Weighted Avg
0	0.92	0.91	0.92	494	0.921	0.92	0.92
1	0.92	0.93	0.92	506	0.921	0.92	0.92

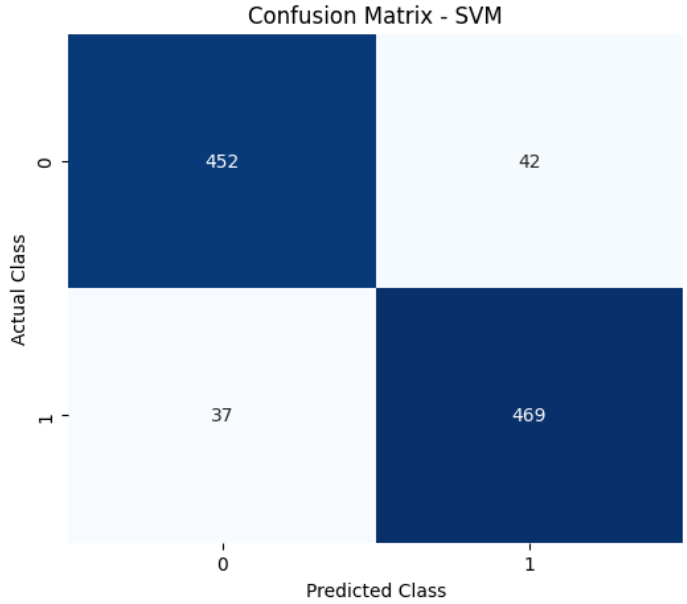


Fig 4.3.6.1 Confusion Matrix of SVM

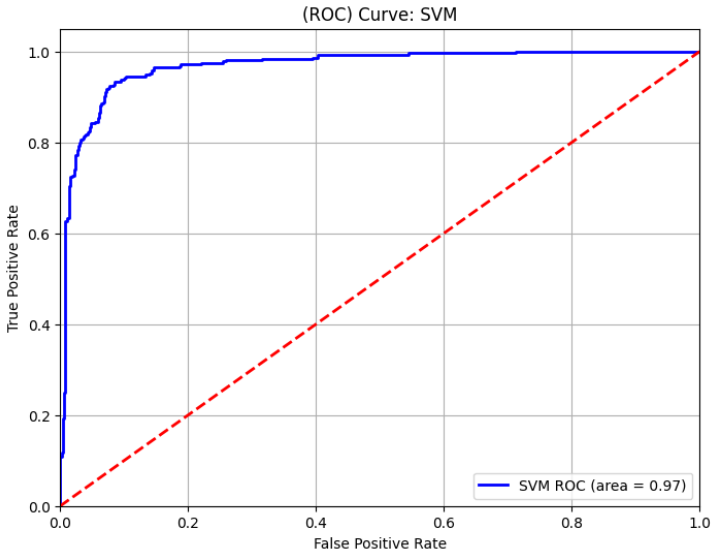


Fig 4.3.6.2 ROC Curve of SVM

4.3.7 KNN

Table 4.3.7.1 Classification Report of KNN

Class	Precision	Recall	F1-Score	Support	Accuracy	Macro Avg	Weighted Avg
0	0.90	1.00	0.95	494	0.944	0.95	0.94
1	1.00	0.89	0.94	506	0.944	0.95	0.94

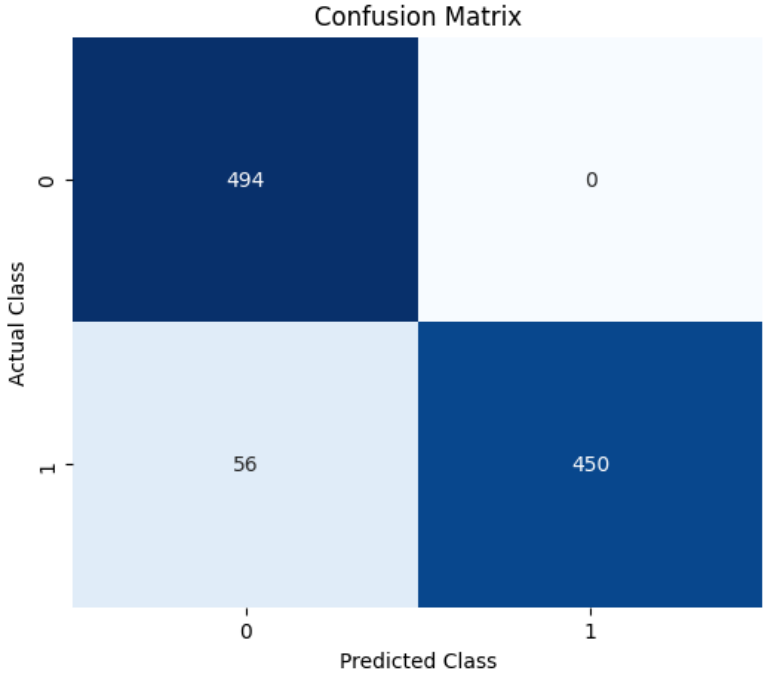


Fig 4.4.7.1 Confusion Matrix of KNN

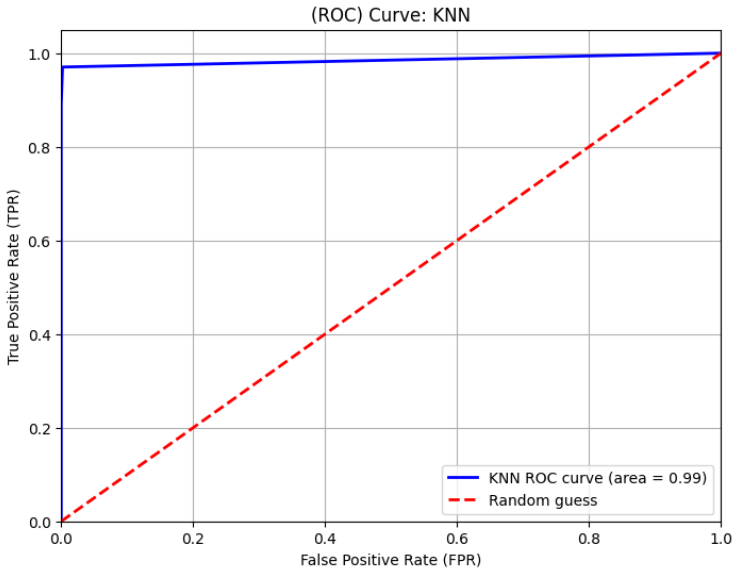


Fig 4.4.7.2 ROC Curve of KNN

4.3.8 MLP Classifier

Table 4.3.8.1 Classification Report of MLP Classifier

Class	Precision	Recall	F1-Score	Support	Accuracy	Macro Avg	Weighted Avg
0	0.96	0.98	0.97	494	0.972	0.97	0.97
1	0.98	0.96	0.97	506	0.972	0.97	0.97

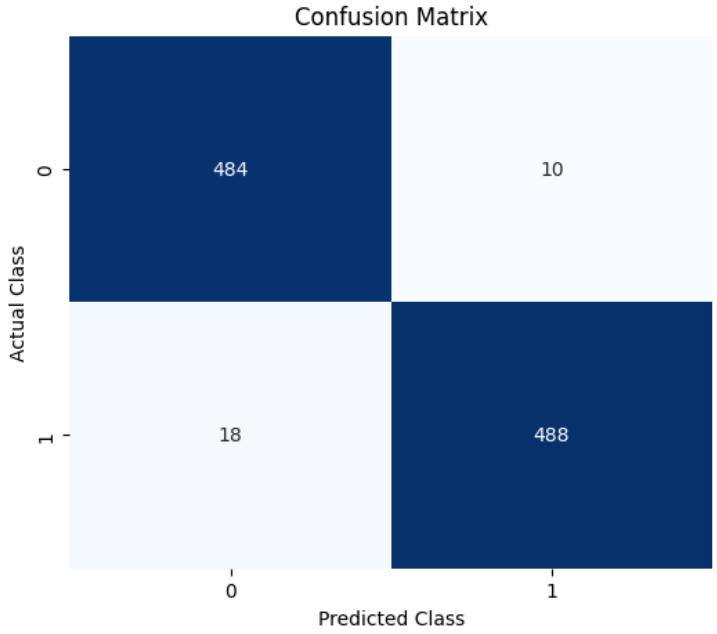


Fig 4.3.8.1 Confusion Matrix of MLP Classifier

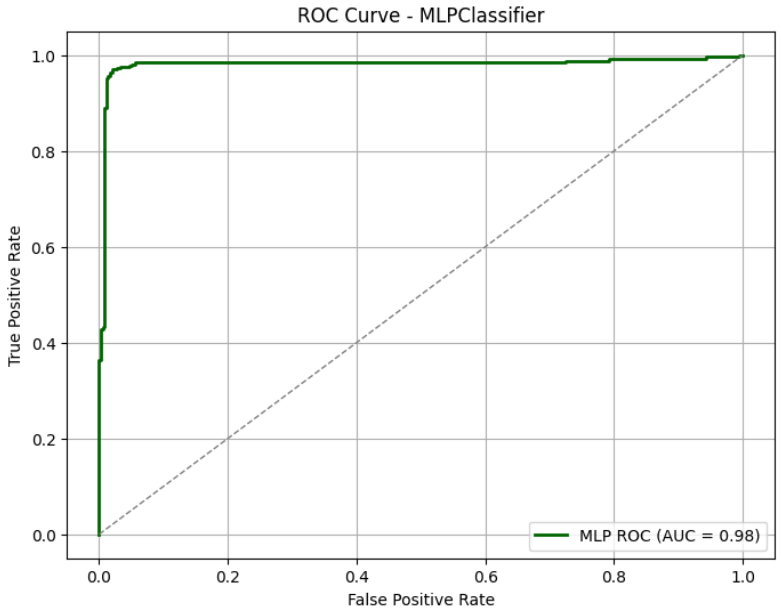


Fig 4.3.8.2 ROC Curve of MLP Classifier

4.3.9 Neural Network

Table 4.3.9.1 Classification Report of Neural Network

Class	Precision	Recall	F1-Score	Support	Accuracy	Macro Avg	Weighted Avg
0	0.98	0.98	0.98	494	0.981	0.98	0.98
1	0.98	0.98	0.98	506	0.981	0.98	0.98

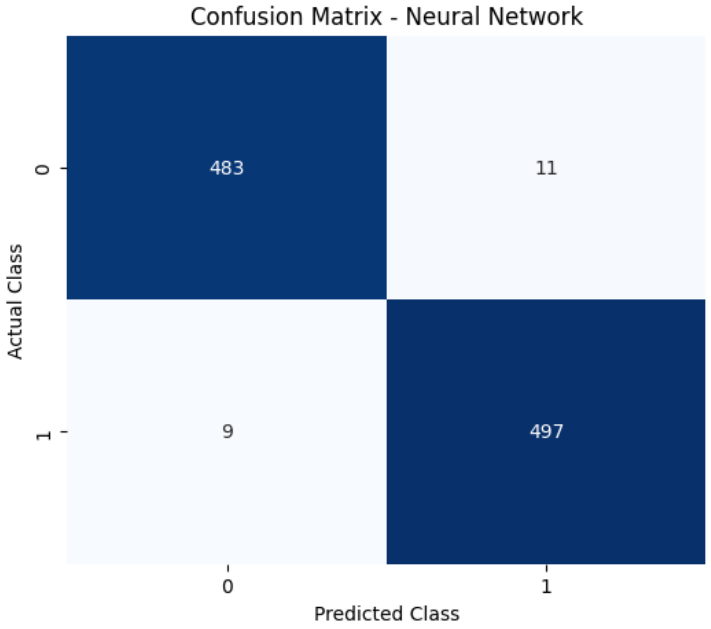


Fig 4.3.9.1 Confusion Matrix of Neural Network

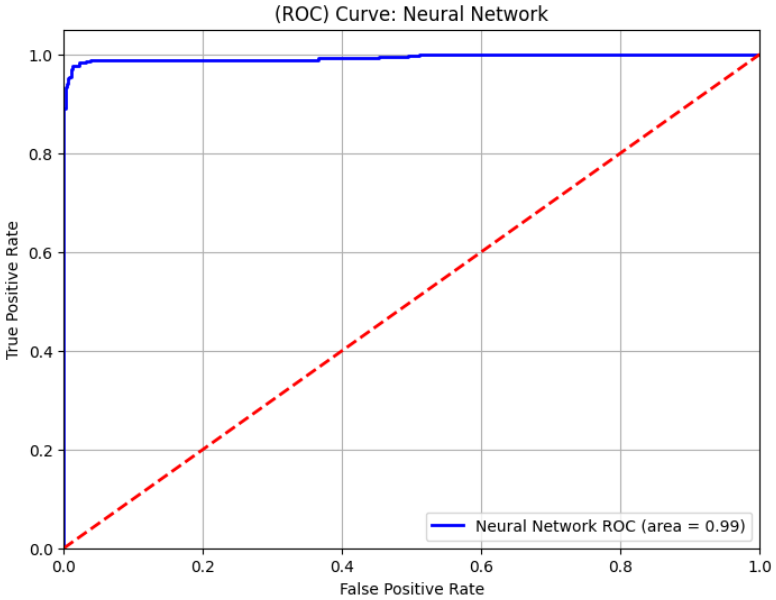


Fig 4.3.9.2 ROC Curve of Neural Network

Table 4.3.1 All Applied Model Accuracy

Model	Decision Tree	Random Forest	Gradient Boosting	XG Boosting	Light GBM	SVM	KNN	MLP Classifier	Neural Network
Accuracy	0.981	0.961	0.916	0.916	0.916	0.916	0.916	0.916	0.916

The accuracy of nine machine learning models was measured on the dataset of IoT security, where the accuracy measures were taken as the most important. Decision Tree and the random forest models recorded the best model with the highest accuracy of 0.981 and 0.961 respectively. The rest of the models such as Gradient Boosting, XG Boosting, Light GBM, SVM, KNN, MLP Classifier and Neural Network all had an accuracy of 0.916. Such findings suggest that tree based models or specifically Decision Tree and Random Forest are better placed to capture the trends in the data than the ensemble and neural network based models. The minor decrease in accuracy of Gradient Boosting, XG Boosting, and Light GBM could be explained by the overfitting or inadequate hyperparameter optimization, whereas the similarity in the results of the SVM, KNN, and neural network models indicate that the data properties are biased towards easier tree-based decision boundaries. On the whole, these results can be used to demonstrate the strength of Decision Tree classifier in detecting internet security threats and emphasize the significance of selecting models that depend on the specifics of the data..

4.3.10 Integrating Web Application

Our site, homepage presents IoT Threat Mitigate; a platform that deals with securing IoT networks against ransomware attacks. It emphasizes machine learning to detect threats in real-time, respond to threats, and monitor threats.

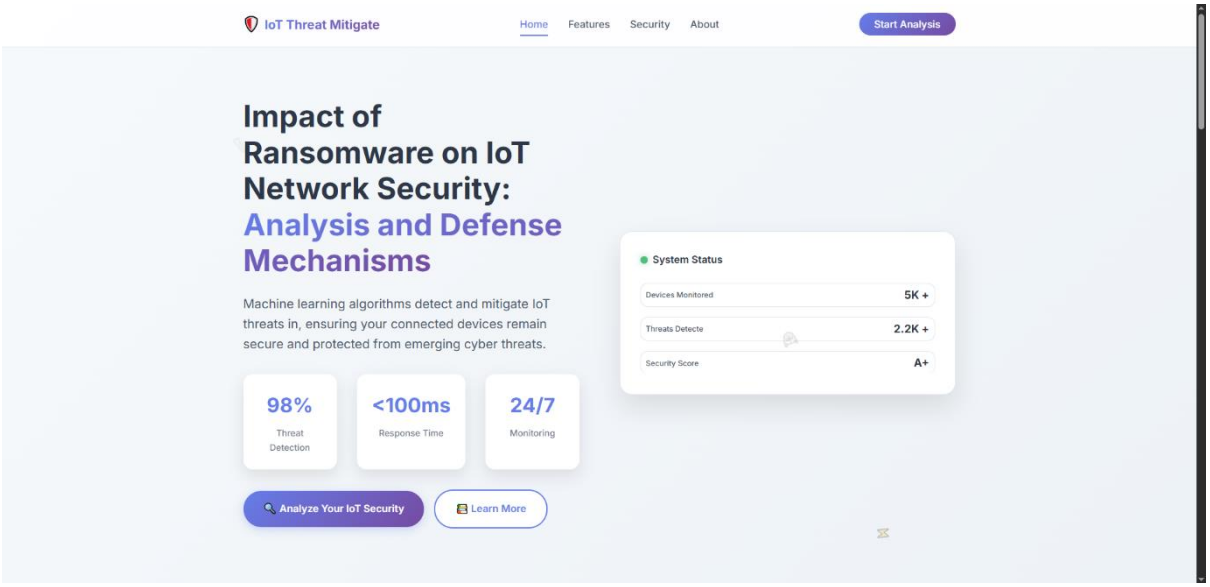


Fig 4.3.10.1 Website Home Page

Threat Assessment of IoT Threat Mitigate allows users to simulate the threat with the ML models and evaluates such factors as data size, processing time, energy consumption, and the type of threat to determine the level of risk.

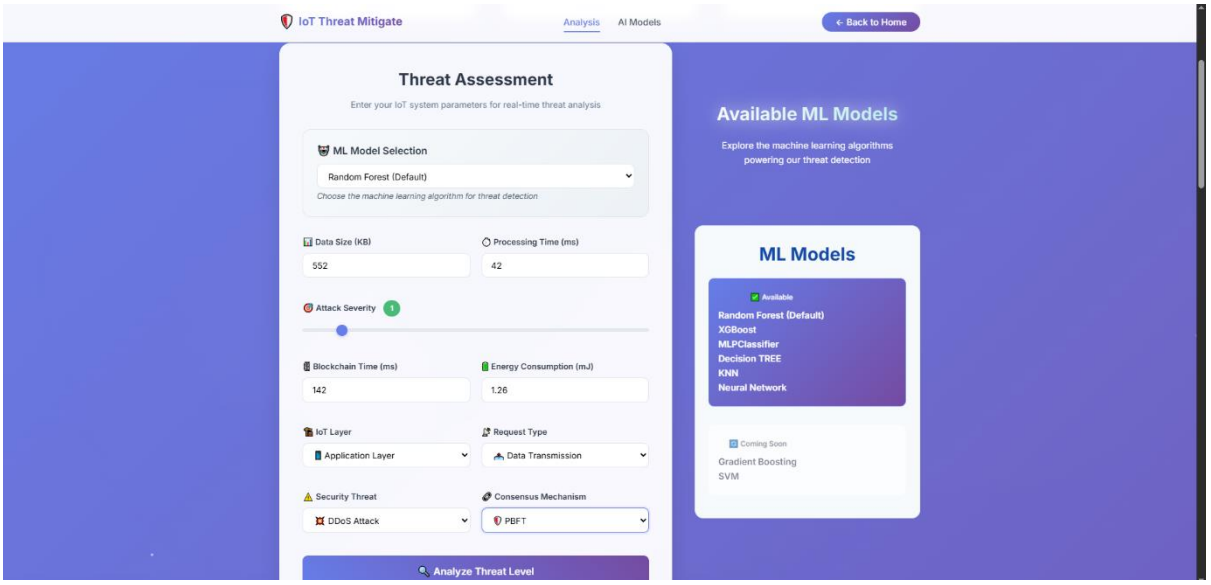


Fig 4.3.10.2 Website Threat Assessment Form

You are 96.7% certain that your IoT system is safe and there is no active threat. The Risk Analysis using the Random Forest indicates a low DDoS threat (severity 1.0) and effective operation of the systems.

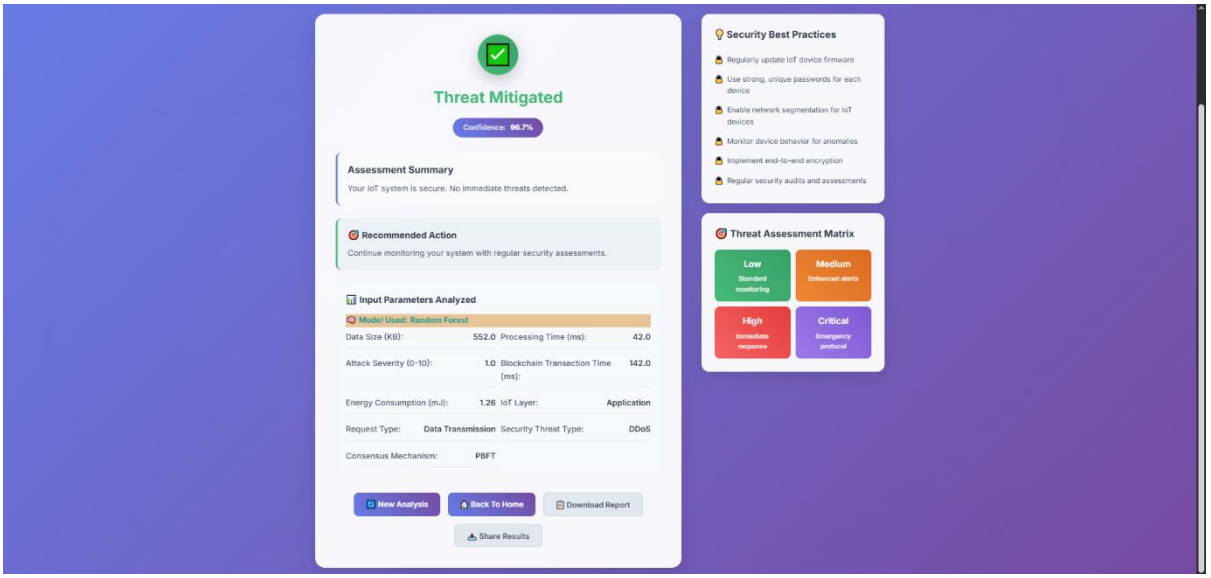


Fig 4.3.10.3 Website Threat Mitigate Result

It advises to check in and adhere to best security practices such as updates on firmware, strong passwords and detecting anomalies.

4.4 Summary

The testing process of nine machine learning models on the preprocessed IoT security dataset proves that the highest accuracy of 0.981 was obtained with the Decision Tree, and then with Random Forest of 0.961. The other models such as Gradient Boosting, XG Boosting, Light GBM, SVM, KNN, MLP Classifier and Neural Network had an accuracy value of 0.916 each. Before training the model, the data was subjected to systematic preprocessing such as label encoding identifiers, one-hot encoding categorical features, and Z-score normalization of numerical features, all features had equal importance. The imbalance between the classes was solved through a combination of SMOTE and Random Oversampling, creating balanced dataset which enhanced generalization and reliability of the model. The findings of the study indicate that tree-based models can effectively obtain complex patterns on the data of the IoT security, although ensemble and neural network models might need additional tuning or more data to attain increased performance. The research highlights that the right and proper model selection and preprocessing is essential to effective and effective IoT threat detection.

Chapter 5

Engineering Standards and Design Challenges

This chapter lists the engineering requirements of the proposed study of the IoT security and explains the design issues encountered in conducting the experiment. The importance of following the established standards so that the research methodology becomes valid, reproducible and reliable is stressed.

5.1 Compliance with the Standards

The use of research standards also guarantees that the procedures and structures are repeatable, assessable, and comparable within research. Here, software, hardware, and communication standards will be pointed out in the research, and alternatives and justifications will be provided.

5.1.1 Software Standards

ISO/IEC 27001 (Information Security Management)

Guarantees the methodical control of security risks in IoT datasets and threat identification experiments.

Alternative: ISO/IEC 12207 (Software Life Cycle Processes), but not so security centric.

IEEE 829 (Software Test Documentation)

Gives systematic records of experiments and model analysis, which makes it reproducible.

Alternative: Non-standardized test procedures; although standardized documents enhance comparisons and consistency with other research.

5.2 Impact on Society, Environment and Sustainability

The protection of the IoT devices directly reflects on the way of everyday life as it ensures personal safety, privacy, and healthcare performance, prevents the data breach, and ensures the continuation of operations in smart homes and other critical systems. In terms of environmental perspective, efficient and secure implementations of IoT can help minimize unnecessary energy use and electronic waste due to device failure and lead to the use of technologies sustainably. At the societal level, the presence of strong IoT security breeds confidence in smart technologies, stable operation of critical infrastructure, and stable and safe provision of services of all industries, which eventually leads to technological development and societal welfare.

5.2.1 Impact on Life

IoT devices are increasingly integrated into daily life including smart homes, healthcare monitoring, industrial automation, and transportation systems day by day. Security breaches in these devices, such as ransomware attacks or data tampering, can have direct and significant consequences on human safety, privacy, and overall quality of life. For example, compromised healthcare sensors could provide incorrect patient data, leading to misdiagnosis or delayed treatment, while attacks on smart home or industrial IoT systems could result in property damage, financial loss, or operational disruption.

This study can help to reduce such risks by paying attention to strong threat detection and prevention systems so that IoT systems would be safe, reliable, and trustworthy. The effective security measures would not only secure the individual users but also increase the trust of people in the society in the overall use of the IoT technologies, which will have beneficial effects in day-to-day life, workplaces, and vital infrastructure.

5.2.2 Impact on Society & Environment

Impact on Society:

The efficiency, automation and convenience of society have significantly risen by the growing rate of use of Internet of Things (IoT) devices in everyday life. IoT has transformed the lifestyle and working experience, including industrial robots and transportation as well as smart homes and medical devices. Nonetheless, this development comes with a high level of danger in the form of ransomware attacks targeting the IoT ecosystem. Such attacks can disrupt everyday activities within the government and citizens and may compromise personal information, and shut down the required services. As an example, a successful ransomware attack on IoT-enabled healthcare devices could threaten patient lives because of their cessation of medical services or altering vital data. The smart cities may be ruined by cyberattacks that can destroy public safety infrastructure, traffic management system, and utilities, leading to chaos and endangering lives. The failure of IoT systems in the industrial environment can lead to revenue losses, breach of production, and threat of workers safety. The primary goals of the study are to identify and analyze and prevent ransomware attacks within Internet of Things environments. The research aims to protect the social frameworks that are increasingly becoming dependent on the Internet of Things by observing the weaknesses and recommendation of effective security measures. It raises awareness among the consumers, developers, and politicians and promotes an active culture of cybersecurity.

Ultimately, a safer IoT setting will ensure a more secure technological implementation, increased trust among the user in the digital system, and further development of the society without the interference of malevolent activities.

Impact on Environment

The Internet of Things (IoT) has dramatically increased in number to the benefit of urban development, manufacturing, energy, and agriculture. The IoT devices are able to improve environmental sustainability by ensuring maximum utilization of natural resources, reducing wastage, and also improving energy efficiency. As an example, energy monitoring devices help businesses to decrease their carbon footprint, and smart irrigation systems in farming help to reduce water waste. Nevertheless, the environment might also have some problems with our increasing dependence on our interconnected gadgets, especially when ransomware attacks are aimed at them. A network of IoT that controls water supply, garbage collection, or power networks could cause a severe damage to the environment in the event of a breach. A ransomware attack could disable smart waste systems that may lead to pollution and unstopped garbage. The IoT equipment used in the energy sector could be hacked and lead to blackouts or overuse of energy which would have a tangible impact on carbon emissions and destruction of ecosystems. This study addresses the problem of internet of things (IoT) cybersecurity that otherwise may surpass the environmental benefits of digital transformation. Powerful mitigation measures to ransomware attacks are established to ensure that intelligent systems remain secure and functional to operate in a productive way without causing a negative impact on the environment. Besides, the research encourages the creation and adoption of IoT solutions that are energy-efficient and environmentally conscious. The resilience of systems to cyberattacks enhances the sustainability of smart technology in the long term and leads to ecological balance. A secure IoT ecosystem is therefore vital regarding technological progress as well as the conservation of the environment.

5.2.3 Ethical Aspects

The work of Internet of Things (IoT) devices in everyday life has potential and moral concerns. One of the principal ethical concerns is user privacy. IoT devices often gather sensitive operational or personal data, and any unauthorized access, especially the one caused by ransomware attacks, may cause the misuse of information and breach of privacy. Consequently, it is important to ensure that there is a guarantee of data security and safe communication protocols. Another ethical factor is transparency. Service providers and manufacturers should specify the method of data collection, storage, and usage. Such openness would help users unintentionally expose themselves to danger. Moreover, the ethical implementation of the IoT systems requires informed consent of users and groups affected by the technology. Another ethical imperative in this research is the need to develop strong and secure systems that will protect users and the society in general. Besides endangering lives, not preventing cyber assaults in the fields that are crucial, such as the healthcare, transportation, or agriculture sectors, also violate moral standards of responsibility and security. Ethics is hence a prerequisite in the creation of a sustainable IoT.

5.2.4 Sustainability Plan

A systematic plan has been developed to ensure the sustainability and effectiveness of this Internet of Things-based solution to ransomware threats. The first step is the implementation of a scaled and secure system design, which can incorporate new technologies and adapt to new needs. To resist the dynamic cyberthreats, the system will be constructed in such a way that it will support fixes and updates without disturbing services. It will be important to carry out regular maintenance and monitoring. This should be done in order to ensure that the system continues operating smoothly through periodic security tests, vulnerability tests, and performance audits. User awareness and training programs will also be implemented to inform the end users on how to handle data, safe practice and how to respond to potential dangers. To reduce expenses without compromising quality, open sources and low costs will be prioritized. In order to exchange finances and information, partnership with government organizations, non-governmental organizations, or educational institutions can be sought. Environmentally, the system will aim at using low energy consumption and use of high-quality hardware to reduce electronic waste. This approach will ensure that the project will be able to have a long-standing effect without compromising the moral or environmental criteria due to the focus on security, effectiveness, affordability, and the environment..

5.3 Project Management and Financial Analysis

This part provides a project management plan, cost assessment, and possible financial aspects of the study on internet of things security and threat identification. Even though it was mostly a research undertaking, knowing the budgetary requirements and options assists in making sure it is viable to go through and the resources are optimized.

5.3.1 Budget Analysis

The estimated budget for the research includes costs for hardware, software, data acquisition, cloud services and miscellaneous resources:

Item	Cost (BDT)	Notes/Alternatives
IoT devices (sensors, hubs)	5,500	Using simulated datasets can reduce cost
Cloud services (storage & computation)	3,300	Local servers are possible, but less scalable
Software licenses & tools (Colab etc)	2,200	Open-source alternatives can reduce cost to near-zero
Data acquisition & preprocessing	1,100	Public datasets can be used
Miscellaneous (power, accessories)	1,100	Required for hardware experiments

Total Estimated Budget: 132,00 BDT
Alternative Budget (Simulations and open-source tools): 44,00–55,00 BDT

Rationale for Choices:

- Using physical IoT devices provides realistic testing but increases cost; simulations reduce expenditure while maintaining experimental validity.

- Cloud services enable scalable computation for model training; local resources are cheaper but may limit large-scale experiments.

- Open-source software minimizes license costs while still supporting full functionality for machine learning and security analysis.

5.3.2 Revenue Model:

Although primarily academic, the research could have potential commercial applications:

- Subscription-based IoT security monitoring services for smart homes or industries.

- Licensing AI-based threat detection models to IoT manufacturers.

- Integration with existing IoT platforms as a security module, generating service fees.

This financial analysis demonstrates that the research is feasible within a moderate budget while allowing flexibility through alternatives, ensuring both experimental validity and practical applicability.

5.4 Complex Engineering Problem

5.4.1 Complex Problem Solving

In this section, provide a mapping with problem solving categories. For each mapping add subsections to put rationale (Use Table 5.1). For P1, you need to put another mapping with Knowledge profile and rational thereof.

Table 5.1: Mapping with Complex Engineering Problem.

EP1 Dept of Knowledge	EP2 Range Of Conflicting Requirements	EP3 Depth of Analysis	EP4 Familiarity of Issues	EP5 Extent of Applicable Codes	EP6 Extent Of Stake-holder Involvement	EP7 Interdependence
✓	✓	✓	✓	✓	✓	✓

Mapping with Knowledge Profile

This section is designed to map the overall problem and EP1 ***(multiple between K3, K4, K5, K6, K8 for attaining EP1)*** to the Knowledge Profile.

Table 5.2: Mapping with knowledge Profile.

K1 Natural Science	K2 Mathematics	K3 Engineering Fundamentals	K4 Specialist Knowledge	K5 Engineering Design	K6 Engineering Practice	K7 Comprehension	K8 Research Literature
✓	✓	✓	✓	✓	✓	✓	✓

5.4.2 Engineering Activities

In this section, provide a mapping with engineering activities. For each mapping add subsections to put rationale (Use Table 5.3).

Mapping with Complex Engineering Activities

This section is designed to map the overall problem and EA's ***(multiple).***

Table 5.3: Mapping with Complex Engineering Activities.

EA1 Range of resources	EA2 Level of Interaction	EA3 Innovation	EA4 Consequences for society and environment	EA5 Familiarity
✓	✓	✓	✓	✓

5.5 Summary

In this chapter, the focus was on identifying and applying engineering standards that ensure scientific rigor, reproducibility, and reliability in the research on IoT security and threat detection. Software standards such as ISO/IEC 27001 were used to systematically manage data security, while IEEE 829 helped structure experiment documentation and maintain clarity in model evaluation. Hardware standards, including IEEE 802.15.4, were adopted to simulate low-power IoT networks realistically, ensuring experiments reflect actual device constraints. Communication standards like MQTT and IEEE 802.3 (Ethernet) were applied to provide scalable and standardized communication frameworks for experimental setups.

The chapter also brings out major design issues being met in the research area. These are the resource constraints of IoT devices, heterogeneous and noisy data, and the requirement of scalable solutions that are generalized to diverse IoT settings, and security-performance trade-offs during the implementation of detection models. To deal with these issues, preprocessing of datasets, balancing and standardization features, and simulation of IoT environments with simple physical devices were needed.

Lastly, the study highlights the effect and importance of secure IoT systems. Securing IoT not only protects the personal safety and privacy, but also helps preserve the environment through minimizing unproductive use of energy and electronic waste and helps the society build trust in new technologies. This chapter illustrates how the research is methodologically sound, and which practical and societal concerns are taken into consideration in the IoT field by balancing compliance with standards with careful design management challenges.

Chapter 6

Conclusion

This paper introduced a machine learning-based infrastructure to identify and address ransomware threats in Internet of Things; a research. The system demonstrated a high level of distinguishing the malicious and benign activities by gathering and preprocessing of the relevant data, addressing the problem of class imbalance, and implementing a variety of classification models. The effectiveness of the approach was confirmed by performance assessment based on precision, recall, F1-score, etc. Although some disadvantages may be mentioned, the suggested solution offers a reasonable basis to improve the security of IoT and suggests meaningful suggestions to improve the implementation in the future by applying real-world practices and learning methods.

6.1 Summary

This project aimed at developing an intelligent and secure Internet of Things which was capable of detecting and preventing ransomware attacks. It began with an all-inclusive understanding of the implications of ransomware on smart devices and networks they are connected to. To address this issue, we collected a relevant set of data sources that are authoritative and thus they represent a variety of ransomware and benign behaviors that are commonly found in IoT space. The preprocessing was done after collection of the data to ensure that it was clean, normalized and coded in order to prepare them to be analyzed. Features have been selectively chosen in order to reduce the computational cost and make the models more fidelity. This was followed by the identification of the best machine learning models to detect the presence of ransomware by applying and analyzing several of them including Decision Tree, Random Forest, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM). Performance indicators such as the accuracy, precision, recall, and F1-score were used as model efficacy. An intuitive user interface and a real time monitoring option made the system much better. This broad approach that considers data collection, data processing, intelligent identification, and real-world application demonstrates a robust solution which can assist cybersecurity in the growing IoT domain.

6.2 Limitation

Despite the promising outcomes of the proposed system in the process of identification of ransomware threats in an IoT environment, a number of limitations can still be noted. To begin with, the training and evaluation dataset were gathered through controlled and trusted sources that might not adequately portray the complexity and diversity of attack cases of ransomware in actual IoT systems. This restricts the generalization on the models in the presence of undetectable or changing attack vectors. Second, some features might not reflect the complete behavioral scenario of the IoT devices despite the preprocessing that made it properly normalized and encoded, which might result to false

positives or false negatives. Also, the more common machine learning algorithms like Decision Tree, Random Forest, KNN, and SVM are very sensitive to feature engineering and might not readily accommodate new forms of ransomwares without regular retraining.

The other weakness is that there is no large-scale real-time deployment and testing. The system was tested in a laboratory environment, and its functionality in the real-life scenario when the IoT networks are exposed to dynamic traffic, a variety of devices, and a constant change in cyber threats is not tested yet. In addition, the models also had computational overhead and latency that was not widely investigated, which can impact the scalability and viability of resource-constrained IoT devices. Finally, the existing system is detection-centered with no automated mitigation procedures, which implies that still the human factor needs to be involved to address the identified threats.

6.3 Future Work

In the future, this project can be expanded in terms of App-based integration allowing one to monitor the project in real-time and identify the danger with the help of an online dashboard easily navigable. Implementation of the system in a real life IoT setup to determine how it performs under real life conditions may result in additional improvements. New types of ransomware may be included to the data, a work system may be developed to automatically respond to eliminate threats quickly, and deep learning may be explored to achieve better accuracy. Such innovations will enhance scalability, usability, and reliability of the system in a number of industries. The other potential option would be automated defensive systems whereby the system is not only able to identify threats, but also act directly, such as disconnecting affected devices or blocking malicious IPs. Finally, the assessment of the system within the real Internet of Things scenarios will also help to verify the system as effective, performant, and adaptable to dynamic conditions.

References

[1] M. Humayun, N. Jhanjhi, A. Alsayat, and V. Ponnusamy, "Internet of things and Ransomware: Evolution, mitigation and prevention," *Egyptian Informatics Journal*, vol. 22, no. 1, pp. 105–117, Mar. 2021. doi:10.1016/j.eij.2020.05.003

[2] R. A. S. Khan and M. N. A. Rahman, "Efficiency of surveillance of TCP packet in IoT in reducing the risk of ransomware attacks," *Journal of Theoretical and Applied Information Technology*, vol. 101, no. 3, pp. 1126–1136, Feb. 2023.

[3] M. Msgna, "Anatomy of attacks on IOT systems: Review of attacks, impacts and countermeasures," *Journal of Surveillance, Security and Safety*, vol. 3, no. 4, pp. 150–73, 2022. doi:10.20517/jsss.2022.07

[4] E. Alotaibi, R. B. Sulaiman, and M. Almaiah, "Assessment of cybersecurity threats and defense mechanisms in wireless sensor networks," *Department of Computer Networks and Communications, King Faisal University*, *School of Computer Science and Technology, Northumbria University*, *King Abdullah II IT School, The University of Jordan*, 2025.

[5] S. Temara, "The Ransomware Epidemic: Recent Cybersecurity Incidents Demystified," *Asian Journal of Advanced Research and Reports*, vol. 18, no. 3, pp. 1–16, 2024, doi: 10.9734/ajarr/2024/v18i3610.

[6] T. R. Reshmi, "Information security breaches due to ransomware attacks - A systematic literature review," *International Journal of Information Management Data Insights*, vol. 1, no. 2, p. 100013, Nov. 2021. doi:10.1016/j.jjimei.2021.100013

[7] Y. L. Khaleel *et al.*, "Network and cybersecurity applications of defense in adversarial attacks: A state-of-the-art using machine learning and Deep Learning Methods," *Journal of Intelligent Systems*, vol. 33, no. 1, Jan. 2024. doi:10.1515/jisys-2024-0153

[8] F. Jimmy, "Cyber security vulnerabilities and remediation through cloud security tools," *Journal of Artificial Intelligence General science (JAIGS) ISSN:3006-4023*, vol. 3, no. 1, pp. 196–233, Apr. 2024. doi:10.60087/jaigs.vol03.issue01.p234

[9] A. Anjum, A. Siddiqua, S. Sabeer, S. Kondapalli, C. Kaur, and K. M. Rafi, "Analysis of Security Threats, Attacks in the Internet of Things," *International Journal of Mechanical Engineering*, vol. 6, no. 3, pp. 2942–[last page if available], Dec. 2021

[10] [1] R. S. Mousa and R. Shehab, "Applying risk analysis for determining threats and countermeasures in workstation domain," *Journal of Cyber Security and Risk Auditing*, vol. 2025, no. 1, pp. 12–21, Jan. 2025. doi:10.63180/jcsra.thestap.2025.1.2

[11] C. S. Yadav *et al.*, "Malware analysis in IOT & Android systems with Defensive Mechanism," *Electronics*, vol. 11, no. 15, p. 2354, Jul. 2022. doi:10.3390/electronics11152354

[12] P. B. Ogini, I. B. Cookey, and C. M. Ajoku, "Impact of ransomware attack trends on

the effectiveness of defence mechanisms: A review of current attack patterns and innovative solutions," *International Journal of Scientific Research in Education*, vol. 18, no. 2, pp. 168–176, 2025.

[13] A. Costin and J. Zaddach, "IoT Malware: Comprehensive Survey, Analysis Framework and Case Studies," vol. –, no. –, pp. –, 2018. [Online]. Available: http://firmware.re/malw

[14] E. Alotaibi, R. B. Sulaiman, and M. Almaiah, "Assessment of cybersecurity threats and defense mechanisms in wireless sensor networks," vol. 2025, no. 1, pp. –, 2025.

[15] S. Saeed, N. Z. Jhanjhi, M. Naqvi, M. Humayun, and S. Ahmed, "Ransomware: A framework for security challenges in internet of things," *2020 2nd International Conference on Computer and Information Sciences (ICCIS)*, pp. 1–6, Oct. 2020. doi:10.1109/iccis49240.2020.9257660

[16] CICIoT2023: "A New Intrusion Detection Dataset for Internet of Things (IoT) Environments," *Data*, vol. 8, no. 1, p. 37, 2023. doi: 10.3390/data8010037.

[17] A. Dabbagh et al., "HoneyIoT: An IoT Honeypot Framework for Exposing Cyber Threats Targeting IoT Devices," in *Proc. ACM WiSec*, 2023, pp. 1–12. doi: 10.1145/3579856.3593538.

[18] Y. Inoue et al., "SPOT NAS Honeypot: Large-Scale IoT Honeypot for Malware Campaign Analysis," *IPSJ Journal of Information Processing*, vol. 32, pp. 68–78, 2024. doi: 10.2197/ipsjjip.32.68.

[19] M. Ashrafuzzaman et al., "Studying IoT Cybersecurity Threats Using Leaked Credentials and Honeypots," *Sensors*, vol. 25, no. 7, p. 2247, 2025. doi: 10.3390/s25072247.

[20] H. E. Al-Rakhami et al., "Federated Intrusion Detection in IoT Using Data Scaling," *Journal of Sensor and Actuator Networks*, vol. 14, no. 1, p. 15, 2025. doi: 10.3390/jsan14010015.

[21] N. C. Luong et al., "Semi-Supervised Federated Learning for Intrusion Detection in IoT Networks," *IEEE Internet of Things Journal*, vol. 10, no. 12, pp. 10320–10334, 2023. doi: 10.1109/JIOT.2023.3269981.

[22] Y. Sharma and B. Bhushan, "Federated Learning-Based Intrusion Detection in IoT Networks," *arXiv preprint*, arXiv:2305.07291, 2023.

[23] J. M. Such and M. N. Kamel, "Adversarial Robustness in Network Intrusion Detection: A Survey," *Electronics*, vol. 14, no. 2, p. 285, 2025. doi: 10.3390/electronics14020285.

[24] R. Qiu et al., "Adversarial Intrusion Detection on NF-ToN-IoT-v2 Dataset," *arXiv preprint*, arXiv:2501.06785, 2025.

[25] A. Souri et al., "TinyIDS: Lightweight Intrusion Detection on Microcontrollers for IoT Devices," *Lecture Notes in Networks and Systems*, Springer, vol. 728, pp. 220–234, 2024. doi: 10.1007/978-3-031-52324-8_17.

[26] A. Abdelaziz et al., "Explainable Deep Learning for IoT Intrusion Detection," *Sensors*, vol. 25, no. 5, p. 1542, 2025. doi: 10.3390/s25051542.

[27] J. Y. Kim et al., "Explainable Artificial Intelligence for IoMT Security Using XAI-XGBoost," *Scientific Reports*, vol. 15, no. 1, p. 2145, 2025. doi: 10.1038/s41598-025-48951-y.

[28] P. Verma et al., "Blockchain-Enabled Deep Learning Framework for IoT Security," *Scientific Reports*, vol. 15, no. 1, p. 1986, 2025. doi: 10.1038/s41598-025-49214-0.

[29] R. K. Singh et al., "XAI-Based Intrusion Detection for Zero-Day Attacks in IoT Networks," *Lecture Notes in Computer Science*, Springer, vol. 14671, pp. 312–327, 2025. doi: 10.1007/978-3-031-64159-1_22.

213-15-4435

ORIGINALITY REPORT

20%	15%	12%	12%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	4%
2	dspace.daffodilvarsity.edu.bd:8080 Internet Source	1%
3	www.mdpi.com Internet Source	<1%
4	Submitted to Melbourne Institute of Technology Student Paper	<1%
5	Submitted to United International University Student Paper	<1%
6	www.preprints.org Internet Source	<1%
7	Shankar Babu, Mahesh Babu Kota. "Synergies in Smart and Virtual Systems using computational intelligence", CRC Press, 2025 Publication	<1%
8	www.ijsre.com.ng Internet Source	<1%
9	arxiv.org Internet Source	<1%
10	www.bright-journal.org Internet Source	<1%
11	www.ijcna.org Internet Source	<1%