SCHOOL OF BUILT ENVIRONMENT, ENGINEERING AND COMPUTING

LEEDS BECKETT UNIVERSITY

THE PERFORMANCE ML AND DL ALGORITHMS FOR INTRUSION DETECTION FOR IOT AND IIOT-BASED SMART ENVIRONMENTS SECURITY USING ENSEMBLE LEARNING IIDS-SIOEL.

Dissertation

Olaleye Cosby Olorunsomo
Student ID: 77263496

Submitted to Leeds Beckett University in partial fulfilment of the requirements for the degree of MSc Data Science

May 2024

Candidate's Declaration

I, Olaleye Cosby Olorunsomo, confirm that this dissertation proposal and the work presented in it are my achievements.
Where I have consulted the published work of others this is always clearly attributed.
Where I have quoted from the work of others the source is always given. Except for such quotations, this dissertation is entirely my work.
I have acknowledged all main sources of help.
I have read and understand the penalties associated with Academic Misconduct.
I have acknowledged all main sources of help.
I have read and understand the penalties associated with Academic Misconduct.

Signed: Olaleye Cosby Olorunsomo

Date: April 2024.

Student ID No: 77263496

THE PERFORMANCE ML AND DL ALGORITHMS FOR INTRUSION DETECTION FOR IOT AND IIOT-BASED SMART ENVIRONMENTS SECURITY USING ENSEMBLE LEARNING IIDS-SIOEL

Abstract

The abstract illustrates the overall description of the research on intrusion detection in IoT and IIoT networks emphasizing key targets, approaches, results, and suggestions. This study will investigate the performance of machine learning (ML) and deep learning (DL) algorithms in detecting intrusions that may occur in the IoT and IIoT networks. Ensemble learning techniques such as stacking and oversampling will be examined for further improvement of detection accuracy. Moreover, the research is committed to measuring the scalability of intrusion detection systems in environments of limited resources and it also assesses its robustness against the dynamic threats in cyberspace.

The methodology includes the descriptive analysis of the database, performance assessments of intrusion detection models using logistic regression, random forest classifier, and convolutional neural network (CNN), review of feature importance, and comparison of model performance. Through the process of identifying the data structure, distribution of the features, and correlation of the variables descriptive analysis helps to bring out these insights. Performance evaluation shows the efficiency of different models to detect intrusions, and the random forest model is shown to be superior in this respect to all other methods. Feature analysis helps to find the most interesting features that determine intrusion detection, consequently, this allows us to understand the patterns of cyber threats.

Additionally, the study finds the best models in the comparison and presents certain areas where the models can be improved, including the integration of anomaly detection techniques and the development of context-aware intrusion detection systems. The research has several constraints such as not having real word deployment and validation which can be overcome in conducting large-scale experiments and field trials.

The results support the relevance of the utilisation of upcoming inventions and adjustable methods as key factors to strengthen the coverage and effectiveness of intruder detection systems across IoT and IIoT networks. As for the next research, the suggested tasks include working on context-aware intrusion detection, bringing together anomaly detection and signature-based approaches, countering adversarial attacks, and using edge computing and blockchains. It should also be considered to create a self-teaching intrusion detection system that can act independently and be capable of self-learning and self-adaptation to dynamic and changing threats.

Table of Contents

Chapter 1: Introduction

1.1 Overview

In a nutshell, this research moves on in the direction of intrusion detection solutions in IoT and IIoT networks, setting a platform for the design of reliable security systems that can be implemented in interconnected environments such as smart factories.

Following the rate of internet of things (IoT) devices' interconnectedness across different areas such as healthcare, manufacturing, transportation and smart cities, the danger of cyber-attacks has been increased. As a result, Machine Learning (ML) and Deep Learning (DL) algorithms become one of the most prominent methods that are applied within the field of intrusion detection to fortify the security of IoT-based and IIoT-based systems.

The rapidly rising need for ML and DL algorithms in smart environment types has been caused by it adaptation to the extremely dynamic and complex data patterns. Using these algorithms, organizations find ways to prevent a security breach and defends sensitive data and critical infrastructure from the attacks end.

The idea of ensemble learning, which takes advantage of an integration of the knowledge of several algorithms, lies at the heart of this process by increasing the detection accuracy and developing the ability to withstand adversarial activities. Thanks to ensemble learning frameworks, like those build on Stacked Integrated Oversampling Ensemble Learning (IIDS-SIOEL), it was possible to achieve quite good outcomes in the security of IoT and IIoT environments.

The principal goal of the investigation is to explore and estimate the capacity of ML and DL algorithms in intrusion detection for IoT and IIoT based smart settings. Through the comprehensive performance metrics and capability analysis of these algorithms, the researchers and practitioners not only identify it capabilities but also its areas of weaknesses and how it could be applied in real life scenarios.

Aiming at this end, the present study will cover the core ideas and methods pertaining to ML and DL techniques. It will illustrate it problem-solving principles for identifying suspicious behaviours and malicious attacks in the IoT and IIoT networks. Additionally, it will focus on the use of the ensemble learning techniques including the stacking and amalgamating of diverse algorithms to reach the synergy and a robust intrusion detection system.

There will be also conducted a comprehensive study of the IoT and IIoT security issues and its specifics within the smart buildings. The fullness of intrusion detection development is complicated by issues like resource limitations, divergent data sources, and constantly shifting attack vectors. Therefore, novel methods that involve ML and DL principles are needed.

By merging the theoretical insights with examples to implement and case studies, the study is aiming at introducing a comprehensive understanding of the state-of-the-art techniques and best practices of using the machine learning and deep learning algorithms for intrusion detection in the cyber-physical world. By providing more meaningful metrics and deeper insights, it attempts to establish the standards and processes closely tied to deploying robust and adaptable security solutions in the emerging space of smart technology ecosystems.

In sum, this research work is dedicated to the accomplishment of the goals in the development of it knowledge and capabilities in saving the IoT and IIoT-based smart environment against cyber-attacks through making the digital ecosystem of tomorrow more reliable, and trustable.

1.2 Statement of the Problem and Rationale

Statement of the Problem:
The intense penetration of Edge into Manufacturing IIoT systems is creating new types of cybersecurity obstacles. Present strategies cannot always cope with security presentations and, as a result, make organizations weak and fall into danger. This work provides a chance to make a comprehensive study and development of the in-depth threat predictions and liabilities of the Edge-IIoT devices in the face of constantly evolving threats in industrial operations.

Rationale:
In the landscape of the Internet of Things (IoT) and Industrial Internet of Things (IIoT), the number of linked devices in IoT and IIoT has introduced a new range of innovations and advancements leading to more effective industries. However, this exponential growth has also brought about a pressing concern: securing smart IoT and IIoT environments. The current cyber systems being utilized by various sectors have become so complexly interrelated that it has exposed it to an array of cyber threat vectors which include data breaches, malware infections, unauthorized access, and control.
The point here is that the cybersecurity in IoT and IIoT environment is a significant issue because of some related factors. Primarily is the big size of the world of connected devices with it various and different sensors, actuators, and embedded systems which cover healthcare, manufacturing, transportation, and energy. This diversity results in certain issues being problematic because each device features a security level that varies and is vulnerable to attacks of different types.
Hence, the resource-stringent nature many of IoT and IIoT devices is the critical factor for deployment of the effective security measures. The things usually have limited capabilities for computing and energy for the traditional security solutions which require complicated crypto algorithms and constant monitoring.
These challenges are complicated further by the volatile and changing nature of cybersecurity threats that IoT and IIoT ecosystems constantly fall vulnerable to. The attack methods are continually changing and coming up with highly advanced methods such as zero-day exploits, polymorphic malware, DDoS attacks, etc. to bypass traditional security safeguards. Furthermore, the tendency to the expanding commercialization of hacking tools and services in the black market has diminished the entry barrier for the aspirant cyber-attackers, resulting in a bigger number and complexity of cybercrime.
The importance of dealing with the security issues at the same time as the internet things and industrial internet deployment spread to critical infrastructure and mission-critical applications cannot be underlined. The effect of a cyber security breach can be so catastrophic. It can result in financial losses, invasion of privacy, or even endanger human lives. This can include smart city and fully autonomous vehicles, as well as industrial control systems and healthcare networks.
Moreover, the IoT and IIoT technologies persevere on growing as well as integrating with the existing infrastructure and obsolete systems, and the problem of security becomes even larger with the increasing size of the attack surface. The long-standing systems, especially,

lead to high risk as it were never built with security features or poorly supported by manufacturers who do not provide updates anymore.

To summarize, cybersecurity in IoT and IIoT-based smart environments evolves into an urgent, complex, and all-important issue that calls for the timely implementation of advanced solutions. Convergence of vast devices, resource depletion, dynamic threat evolution as well as critical applications makes the urgent issue of implementation of well-built intrusion detection system plugged-in with advanced ML and DL algorithms such as those that are investigated in this research. When the researchers and the practitioners deal with all the mentioned challenges, it will help secure IoT and IIoT ecosystems and make sure that the integrity, reliability, and resilience of the digital infrastructure of smart environments which is going to be built in the future are guaranteed.

1.3 Aim and Objectives

Aims

Main goal, that is to strengthen the cybersecurity infrastructure of IoT and IIoT for and where ML and DL algorithms are used for intrusion detection to overcome the vulnerabilities and protective systems in an interconnected environment.

Objectives

- To evaluate the performance of ML and DL algorithms in detecting intrusions within IoT and IIoT networks.
- To investigate the effectiveness of ensemble learning techniques, such as stacking and integrated oversampling, in improving intrusion detection accuracy.
- To assess the scalability and resource efficiency of intrusion detection systems deployed in resource constrained IoT and IIoT environments.
- To analyse the robustness of intrusion detection mechanisms against evolving cyber threats and attack vectors targeting smart environments.
- To provide insights and recommendations for the design and implementation of resilient security solutions in IoT and IIoT ecosystems.

1.4 Outline

The structure of the research, which evaluates the aspects of ML and DL algorithms utilizing ML and DL algorithms for intrusion detection in IoT and IIoT-based smart environments, is an outline that is based on a systematic aim to study, analyse, and present findings.

The 1st chapter is an introduction that leads the study by presenting the rationale for cybersecurity in the interconnected huge organization. by means of this, the objectives of research will be laid down to enhance the security via ML and DL methods being utilized for intrusion detection. Further, it has an outline of the rationale for the study, which explains the greatest problem and the obstacles encountered while ensuring a smart environment.

Chapter 2 presents the literature review and providing a comprehensive review of the current research landscape, frameworks, and techniques for intrusion detection in IoT and IIoT systems. The section challenges these algorithms' ML and DL performance, ensembles' learning algorithms, and intrusion detection systems that perform similar tasks. Defining the lenses through which previous research can be synthesized and analysed, this chapter builds on the theoretical framework and informs the research methodology.

Chapter 3 enumerates methods used in the research, involving experimental setup, data collection, and evaluation metrics to check the intrusion detection performance of ML and DL algorithms. In this section, it describes the stages of ensemble learning technique evaluation to figure out if stacking and integrated forms of oversampling increase the effectiveness of detection and help to resist cyber threats.
Chapter 4 on the design and deployment of IoT and IIoT-based smart environments-specific intrusion detection systems is the topic. This chapter details the architectural framework, algorithms, and integration techniques that the system employs to create highly resilient, scalable security solutions. It details resource optimization, real-time monitoring, and adaptive learning schemes necessary for the successful purpose of putting a system in different and dynamic environments.

Chapter 5 offers the findings and discussion of the study which includes the empirical analysis and experimentation conducted in the study. It demonstrates these performance measures, compares these algorithms, and provides insights from the exercise of the evaluation of ML and DL algorithms in intrusion detection scenarios. This part of the article encourages a more advanced comprehension of the main benefits, weaknesses, and actual implementation of the suggested protective measures.
Chapter 6 focuses on a summarization of the findings and implications that have been uncovered, presenting it in a coherent way and drawing up a complete conclusion. It stresses the consequences, the directions, as well as recommendations, which should serve as a guild for further research and practical applications.
In conclusion, Chapter 7 represents the summary of the research findings mentioning its key contributions, as well as limitations. The paper provides a special consideration for academia, industry, and politicians as it identifies the areas of further research, technological innovation, and collaboration on the Internet of Things (IoT) and Industrial Internet of Things (IIoT) ecosystems.

Chapter 2: Literature Review

2.1 Introduction

The literature review is what anchors the research landscape by providing an all-encompassing summary of the existing knowledge regarding the frameworks, methodologies, and concepts associated with the those of the IoT and IIoT-based smart environments through the application of ML and DL models for intrusion detection.
Interconnected devices in IoT and IIoT networks are one of the drivers for the evolution of many industries thanks to new innovations and efficiency. Subsequently, this exponential development has also made the space environment vulnerable to the growing threat of cybersecurity attacks that demand exceptional defence measures to protect critical assets and infrastructure. A vast amount of research concentrated on the capability of different algorithms, including Support Vector Machines (SVM), Random Forests, Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) (Hindy *et al.* 2020). for identifying intrusions and anomalies in network traffic, sensor data, and system logs. In the past, this research has demonstrated interesting results by showing that both ML and DL approaches can surpass the traditional rule-based methods and be able to recognize new and evolving attack vectors (Baniasadi *et al.* 2022).

In addition, ensemble learning approaches are a greatly exploited way to augment the detection accuracy and stability of IoT and IIoT systems. When various base learners such as decision trees, neural networks, and SVMs are typically combined in methods of the ensemble, such as bagging, boosting, and stacking, remarkable efficiency in misclassification reduction, dealing with imbalanced data, and boosting the detection performance have been reported. The SIOEL technique which merges ensemble classifiers with oversampling strategies is a novel approach; thus, it is the one that addresses the class imbalance issue and improves notable class detection.
Moreover, literature review sheds light on the unique intricacies and factors in safeguarding both the IoT and IIoT ecosystems. Limitations of resources, highly heterogeneous data sources, changing network patterns, and expanding reconnaissance range present a range of challenges when building a smart environment's intrusion detection mechanisms. The resolution of the above-mentioned challenges necessitates the entire system's understanding of the technologies, protocols, architectures, creative approaches embedded in the ML and DL paradigms, and innovative systems.
Finally, the literature review gives a detailed picture of currently deployed enhanced methods, tools, and best practices in advanced ML and DL algorithms for intrusion detection in IoT and IIoT environments. In this section, the present study builds on prior research by synthesizing the findings and critically evaluating them. This section sets the foundation for the present research and thus informs the research methodology, experimental design, and theoretical framework. Exploiting the findings from the already existing research study in this research venture one aims to give knowledge and capability in securing IIoT and IoT-based smart environments from cyber-attacks thereby building trust, reliability, and resilience in the digital ecosystem.

2.2 Exploring the Efficacy of Machine Learning and Deep Learning Algorithms in Intrusion Detection for IoT and IIoT Networks

In the universe of the Internet of Things (IoT) and the Industrial Internet of Things (IIoT), the spread of interconnected devices has resulted in amazing chances through several realms. This, nonetheless, contributes to the development of comprehensive cybersecurity environments that can withstand emerging cyber threats. The center of this task is the assessment of ML and DL algorithms to identify intrusions in IoT as well as IIoT devices. This topic focuses on the efficiency of diverse machine learning and deep learning methods, which examine it performance, perspectives, and limitations as tools for raising the safety level of interconnected systems (Soliman *et al.* 2023).

A formal assessment of ML and DL algorithms is the most powerful weapon in the battle to strengthen the IoT and IIoT's network security. Such algorithms provide prospects of flexible intrusion detection mechanisms, which can detect abnormal patterns from massive and diverse data streams. Rigorously researchers lead to the exposition of the suitability of different algorithms which clear view of their detection precision, computational efficiency, and robustness against malicious behaviours (Ge *et al.* 2021).

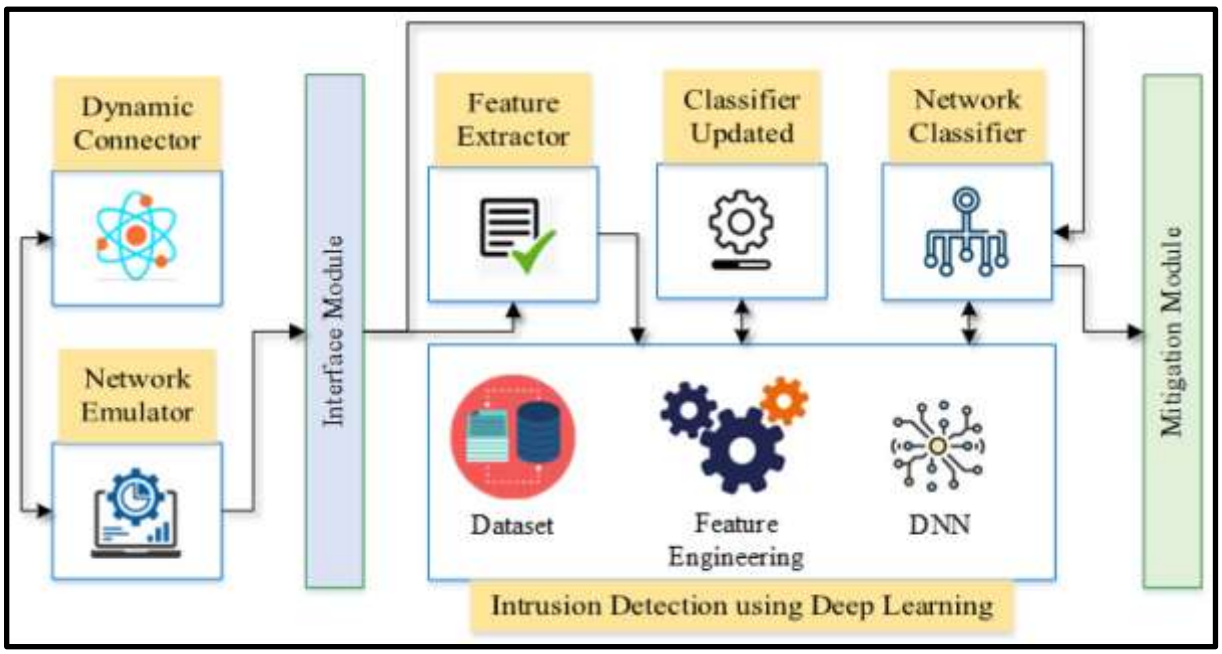


Figure 1: Efficacy of Machine Learning and Deep Learning Algorithms in Intrusion Detection

(Source: mdpi.com, 2024)

The evaluation of the ML and DL algorithms in intrusion detection is characterized by such diversity of techniques and methodologies that are being implemented. Algorithms constitute a wide spectrum that includes traditional classifiers for instance Support Vector Machine (SVM) and Random Forest to neural network systems such as Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN). Every algorithm brings only some benefits and

poses some problems too therefore, when choosing the algorithms for any IoT or IIoT the comprehensive analysis is needed.

Moreover, the test is not confined to just proficiency of algorithm but includes practical aspects such as scalability, resource usage and real-life applicability. Researchers examine the learning algorithm complexity when deploying ML and DL algorithms in the devices that are in limited resources, and while investigating the optimization strategies to decrease the learning algorithm complexity without giving up the detection accuracy. In addition, the researchers analyse the scalability knack of intrusion detection systems, checking it capacity to grow with diverse and changing network topologies along with the emergence of fresh and complicated attack vectors.

Moreover, the assessment of ML and DL algorithms is concerned not only with metrics, but also with the comparison to firmly established datasets and baselines. Researchers make use of datasets that are freely available to the public, such as the NSL-KDD and CICIDS2017 datasets, to evaluate the efficiency of the algorithms as well as generalization in a wide range of network traffic patterns and attacks. To obtain new technics and high-performance standard, researchers having comparison between old methods and modern practices.

Additionally, the assessment of ML and DL algorithms involved in intrusion detection is not confined to only standalone approaches but also includes the ensemble learning method. Ensemble methods which include bagging, boosting, and stacking looks at the collective wisdom of the many base learners in order to improve data detection, validity and resilience against fake positives and negatives (Hamouda *et al.* 2021).

2.3 Exploring Ensemble Learning Techniques for Enhanced Intrusion Detection in IoT and IIoT Networks

This theme revolves around investigating ensemble learning algorithms, which consist of aggregation or bagging (stacking) and integrated oversampling, aiming to highlight it contribution to the security of interdependent systems (Kasongo, 2021).

Combination learning approaches deliver a powerful tool capable of enhanced detection of intrusion via accumulation of base learner Intel. One of the primaries focuses of the empirical study is to examine stacking, a meta-learning algorithm that stacks predictions from different classifiers together to make a final prediction (Gad *et al.* 2021).

Beyond the basic formation and participation in the ensembles, the investigation explores new techniques and hybrid frameworks which support the intrusion detection precision (Nandanwar and Kataria, 2024). The study examines the merging of ensemble learning approaches with anomaly detection systems, anomaly ensembles, and hybrid architectures featuring the domain knowledge and experts' systems. By combining, the unique advantages of different techniques, researchers try to strengthen the power of IOT and IIOT networks intrusion detection for more resilient and reliable interconnected systems this is the way (Smys *et al.* 2020).

Evaluation of ensemble learning methods for advanced intrusion detection in both IoT and IIoT is an overarching and multifaceted task comprising search for new algorithms, optimisation, and practicality. Through the specification of stacking and oversampling methods integration, the specialists aim to improve the state of technology thus increasing the level of cybersecurity and making the intricately connected ecosystems the resistant systems of tomorrow.

2.4 Scalability and Resource Efficiency in Intrusion Detection for Resource-Constrained IoT and IIoT Environments

The expansiveness of the Internet of Things (IoT) and Industrial Internet of Things (IIoT) introduces the scalability and resource effectiveness of intrusion detection systems as one of the critical factors to safeguard the security of interconnected environments. This theme investigates the evaluation of intrusion detection system scalability and resource efficiency in IoT and IIoT environments where computational resources are limited, which focuses on pinpointing the obstacles, strategies, and best practices for the effective implementation of security measures in those circumstances (Abdel-Basset *et al.* 2020).

Facing a variety of IoT and IIoT devices, when sometimes limited by hardware and software resources, intrusion detection becomes more complicated than ever. Therefore, researchers responsible for the assessment of intrusion detection including data volume, variety, and velocity produced by connected devices are to be considered. Scalability is an attribute for intrusion detection system, which supports its capacity to effectively handle and analyse sizable amount of network traffic, sensor data and system logs, ensuring detection accuracy and responsiveness under dynamic network conditions (Khraisat and Alazab, 2021).

The target of the study is to search for lightweight and distributed intrusion detection architectures that are suitable for IoT and IIoT systems characterized by limited resources. Scientists explore new ways of distributed detection, such as edge computing ideas, decentralized solutions, and federated learning models. The intention is to allocate less computational load to central servers and distribute the tasks amongst multiple interconnected devices. The researchers aim to increase scalability and reduce latency employed in the intrusion detection systems adapted to the realistic IoT deployments by means of the distributed computing resources and edge intelligence.

In this area, the assessment covers the topic of scalability, but also resource efficiency, which considers the energy consumption, memory footprint, and processing overhead. The resource poor IoT devices, which are commonly equipped with a low computational power and a battery, require lightweight detection algorithms and optimization strategies to minimize energy consumption and to facilitate longer device lifespan. Scientists develop methods of model compression or feature selection among other hardware acceleration features to ensure the utilization of resources and the enhancement of energy efficiency in intrusion detection systems deployed for both IoT and IIoT (Nuaimi *et al.* 2023).

On the other side, besides algorithmic optimizations, the research tries to investigate the influence of network communication overhead on intrusion detection system scalability and resource efficiency. The existence of a lot of connected devices and various communication protocols forces the network through increased transmission data overheads, packet processing overheads, and network bandwidth resources. Researchers make it investigation on the strategies that mitigate the maintenance overhead, like data aggregating, protocol optimization and adaptive sampling so that there is no network congestion in a distributed intrusion detection system and scalability is on the higher side.

Whilst doing that, the article talks about the interconnection between scalability, resource efficiency, and detection accuracy in the Intrusion Detection Systems implemented in the resource-constrained Internet of Things and Industry 4.0 environments. Distributing the load by separating functions is an effective tactic that can help improve performance. Organizing

the application this way can reduce dependency on a single computer and provide more fault tolerance.

Furthermore, researchers focus on the applications of machine learning and deep learning algorithms in the development of intrusion detection models that can scale and improve resource efficiency. To prevent costly computations and memory footprint while maintaining detection accuracy, researchers plan to use lightweight models with efficient frameworks. Moreover, researchers focus on approaches including transfer learning, online learning, adaptation of current models to environment changes, and dynamic environment characteristics to empower intrusion detection models on the Internet of Things and industrial Internet of things scenarios (Mohy-eddine *et al.* 2023).

In general, the study of scalability and resourcefulness of intrusion detection systems in constrained IoT and IIoT environments consists of explorations in algorithms, designs of system architecture and practical considerations. Through exploration of obstacles and approaches to greater scalability and resource efficiency, researchers aim to move the existing cybersecurity state-of-the-art while helping to realize dependability and strength of interdependent ecosystems where resources are limited.

2.5 Ensuring Robustness in Intrusion Detection Mechanisms Against Evolving Cyber Threats in Smart Environments

In the changing and interactive smart environment, it is a problem to be sure against the continuous cyber threats. This theme is debating the high resistance of intrusion detection techniques to changing cyber threats and penetration ways which are aimed at smart surroundings. It seeks to discover resiliency, adaptability, and the efficiency of intrusion detection systems through research so that professionals, as well as individuals, have enough practical knowledge to mitigate emerging hazards and to maintain the security of the network ecosystems (Campos *et al.* 2022).

A large variety of ways that smart environments can be attacked by hackers themselves, with the techniques ranging from network-based attacks to physical tampering and social engineering exploits (Sarhan *et al.* 2022). Consequently, threat identification process should be more dynamic and encompass more holistic landscape of emerging threats like ransomware, IoT botnets or supply chain attacks. Investigators seek to uncover the techniques, tricks, and procedures applied by adversaries thus educate developers to build and implement real-time detection systems that can detect and counter emerging threats (Alsaedi *et al.* 2020).

The adaptability and responsiveness of intrusion detection processes to the changing attack actions and tactics will be one of the main issues of this study. Researchers do the job to analyse the pros and cons of signature-based detection, anomaly detection, and behavioural analysis techniques involved in different IoT and IIoT environments to detect both known and unknown threats. Such benchmarking helps to evaluate the level of efficiency, applicability, and adaptability with respect to the ongoing and fresh hacking efforts.

Besides that, the analysis covers not only the intrusion detection in a shot but the resilience and survivability of intrusion detection systems in the face of advanced and persistent attacks. Scientists are attempting to improve intrusion tolerance techniques, implementing diversity-based defences, and applying deception methods to make intrusion detection systems more resilient to evasion tactics and zero-day exploits. Through the implementation of redundancy,

diversity, and obfuscation, researchers attempt to defeat attackers' methods of getting away undetected and breaching the smart environments.

In addition, the study scrutinizes the influence of system context factors, namely network topology, system architecture, and user conduct on the robustness of intrusion detection techniques. The investigation focuses on the communication among the environment, sensor data, and network traffic flow to detect an abnormal behaviour that might be a signal of possible security violation. Through the use case of context information and domain knowledge, the researchers want to make the intrusion detection system in smart environments more precise and reliable as well as be able to decrease the false positive and negative rate (Lakshmanna *et al.* 2022).

Alongside algorithmic principles, the role of threat intelligence, threat modelling and information sharing are also studied in this regard with a view to upgrading the robustness of ID mechanisms. With the help of emerging threats, vulnerabilities, and IOC (indicators of compromise) the real-time threat intelligence feeds, the researchers are trying to enrich the intrusion detection systems with the updated information. Also, researcher teams examine ways of collaboration like sharing information and threat hunting which contribute to the enhancement of situation awareness and future threat detection.

To sum up, measuring the stability of intrusion detection systems against the changing cyber threats in smart environments is a complex project involving intelligence on threats, algorithmic modelling, and contextual sensemaking. By revealing strategies or approaches for improving intrusion detection resilience and thus advancing cybersecurity the researchers contribute to the state-of-the-art in cybersecurity, giving reliability and resilience in interconnected systems as well as in evolutionary cyber threats.

2.6 Designing Resilient Security Solutions for IoT and IIoT Ecosystems: Insights and Recommendations

In a context where Things on the Internet of Things and Industrial Internet of Things are evolving rapidly, making sure the security and robustness of the interconnected ecosystems is essential. This topic explores information on the approaches and guidelines for IoT and IIoT security solutions deployment with the ability to withstand shocks. Through the integration of existing knowledge, models, and best practices as well, researchers will be able to shed light on both strategies and approaches that may strengthen the physical security of both, electronic and digital systems, against newly emerging threats and vulnerabilities (Nimbalkar and Kshirsagar, 2021).

The development and implementation of robust security solutions against IoT and IIoT ecosystems necessitate a complete understanding of the various challenges and concerns specific to interconnected systems. Researchers investigate the IoT/IIoT deployment protocols including architecture, communication routines, and deployment scenarios for detection of the possible security flaws and attack vectors. Researchers, by its nature, work to uncover the threat landscape and attack surface to inform the development of security solutions that effectively address diverse cyber threats and guarantee the accuracy of critical systems.

The main investigation idea is studying of security-by-design principles and ways to implement it into the lifecycle of designing and developing IoT and IIoT solutions. Researchers promote the use of a proactive approach rather than reactive practice, with a specific focus

on threat modelling, risk assessment, and security requirements elicitation as the first tasks to be performed. Through the inclusion of security concerns during system architecture, data modelling, and communication protocols, researcher intends to minimize the risk of security violations during the system design stage as implemented in the actual system, ultimately reducing the odds of exploitation (Friha *et al.* 2023).

In addition, the design of robust defense strategies requires the implementation of defense-in-depth architectures and multi-layered controls that could reduce the impact of malicious attacks. Experts of information security look for a composite approach which may consist of network segmentation, access controls, encryption, intrusion detection, and anomaly detection methods. The researchers design multi-defense layers to create redundancy and resilience in cyber systems. It means that attackers fail to find and exploit such single points of weakness or inadequacies in the strategy of security managers.

On the other hand, the implementation of reliable security solutions in IoT and IIoT ecosystem can come from a holistic approach that considers all factors like technical ones, organizational, regulatory, and societal ones. Scholars are working out workable governance models, compliance plans, and regulations that pertains to the deployment of IoT and IIoT to make sure that it harmonises with the applicable standards and good practices in the industry. In addition, researchers explore the role of stakeholders consisting of manufacturers, developers, operators, and users and the role it plays in creating a security culture while collaborating across the IoT ecosystem.

Besides that, the researchers widely examine the part of the emerging technologies and innovations in building up the resilience of digital security for IoT and IIoT. Scientists review the capacity of blockchain technology, decentralized identity management and secure hardware modules in assuring protective authentication, data integrity and secure communication networks that are tamper-proof in interconnected systems. With modern technologies utilization, the researchers strive to face IoT and IIoT security issues and make these networks more resilience against the ever-growing cyber threats.

The last statement reflects the comprehensive nature of providing insights and recommendations for the resilient security solutions designing and implementing in the IoT and IIoT environments composed by the technical, organizational, and social aspects (Tsimenidis *et al.* 2022). Researchers are pursuing this goal by integrating knowledge, models, and standards to arrive at robust and reliable security techniques that can neutralize cyber risks, buttress the integrity, dependability, and reliability of networked systems amid digitalization.

2.7 Literature Gap

The exploration of intrusion detection in both IoT and IIoT systems brings out an existing literature gap on the aspects of integrating context information and domain-specific knowledge into the intrusion detection mechanisms. While there is profound research on the algorithm's performance and techno-optimization, there exists an obvious deficit of such execution that consider the elements of context and environmental dynamics in the smart environments.

Within this context, the concern is that the literature is parsimonious in the exploration of the utilization of contextual information such as network topology, system architecture, user

behaviour, and environmental context to make intrusion detection systems in IoT and IIoT environments more efficient, reliable, and resilient. As part of the contextual awareness involved in intrusion detection, researchers must be able to differentiate the normal behaviour patterns from anomalous actions, and as an outcome of it, erroneous positives and negatives will be reduced, and detection accuracy will be enhanced.

Further, the literature gap relates to incorporation of domain specific knowledge, e.g., of industry regulations of standards, and requirements of operation into intrusion detection frameworks that are ready for deployment in the scenarios of IoT and IIoT. Existing studies concentrate on algorithmic and generic datasets technique but it lack researches which is along the line in IoT application domains such as health care, manufacturing, transport and smart cities.

The lack of such literature creates a big gap which, however, is crucial for further development of intrusion detection technology for IoT and IIoT surroundings. Through factoring into intrusion detection devices details information along with domain knowledge, researchers will be able to develop thinking security mechanisms that are adaptive, robust, and effective solutions for mitigation of emerging threats and ensuring the integrity and dependability of smart environment systems. The upcoming research projects should focus on interdisciplinary co-working, data sharing and real-world validation to overcome the gap in IoT security literature. It defines the field for the future linked by bridges.

2.8 Theoretical Framework

Quite on the sides of developing and improving the intrusion detection mechanisms for the IoT and IIoT system, establishing a solid theoretical foundation is essential. Here, the theoretical foundations on which the study of intrusion detection in interconnected systems rest, are addressed, which includes basic concepts from the field of cybersecurity, machine learning and network security.

The cybersecurity resilience lies at the center of the theory framework, and it means that the systems have the ability of predicting, enduring, and coping with cyber-attacks and security breaches. Applying resilience theory, researchers attempt to explain factors that form the effective defence response of intrusion detection mechanisms for IoT and IIoT environments. Through studying the various factors that may impact resilience, the researchers can develop and implement novel intrusion detection systems with the ability to address ever changing threats and remain functional even in the face of possible vulnerabilities.

Also, the theoretical framework includes the principles from machine learning and data mining fields, which are building block machine learning algorithms and techniques. The researchers work on the fundamentals of supervised learning, unsupervised learning, and semi-supervised learning, that enable the models to be taught on the labelled and unlabelled data. Using machine learning algorithms, like decision tree, support vector machines, and neural networks researchers aim to discover patterns and anomalies which are related malicious activities on the network traffic and the sensor data.

In addition, it comprises anomaly detection principles, which act as a cornerstone in the detection of new and unknown threats not seen before. Researchers go into not only the mathematical backgrounds of anomaly detection techniques but also the statistical analysis, clustering algorithms, and outliers' detection methods. By focusing on anomaly detection

methods, researchers hope to find any deviations in the normal behaviour patterns and detect security breaches in the process being carried out.
Moreover, the framework makes use of machine learning and anomaly detection. Network security and cryptography are the two branches more crucial for making sure that data is kept confidential, intact, and accessible. Researchers study cryptographic algorithms, protocols and digital signatures which are the main basis for encryption and protection of communications and exchanged data in IoT and IIoT systems. Researchers try to achieve the main objective of securing sensitive information and decreasing the possibility of eavesdropping, tampering and data manipulation in transit with the help of cryptographic techniques.
Also, theoretical framework encompasses the principles of the information theory and the signal processing, which are applied in isolation of sensor data and network traffic analysis in IDS. Scientists conduct studies related to entropy, signal-to-noise ratio, and feature extraction, meaning it can analyse the data patterns that help in identifying any anomaly. Using information-theoretic parameters and signal processing techniques researchers would empower intrusion detection models for IoT and IIoT by gaining higher discriminative ability and accuracy.
Also, the theory is built on the principles obtained from system theory and control theory disciplines, which bring in the understanding of the process dynamics and feedback formed in complex cyber-physical system. Researchers examined the ideas of feedback loops, control policies, and system dynamics which are under the control of intrusion detection tools to withstand changes of environmental conditions and cyber threats. To that end, researchers use an approach that is system-theoretic to develop and implement robust IDSs that behave well in terms of stability and functionality while there is uncertainty and disturbances in smart environments.
In the end, the theoretical framework will provide the reader with the complete foundation to understand and explore intrusion detection mechanisms in IoT and IIoT environments. Security researchers seek to leverage the core concepts from cybersecurity, machine learning, network security, cryptography, information theory, signal processing, and system theory to ensure that it is knowledgeable about the many issues as well as the prospects associated with securing system integration against the ever-emerging cyber security threats. Through the combination of theoretical notions with practical evidence gathered from the demonstration and case studies, researchers could innovate the existing state of the element and thus improve resilience, reliability, and security for Internet users of the future.

2.9 Summary

The literature review given has a multi-dimensional exploration of intrusion detection for IoT and IIoT environments, and this is to provide an exhaustive understanding of the current best practices, challenges as well as opportunities in securing the ever-growing interconnected world from new cyber threats. Researchers have taken the interdisciplinary lens spanning computer security, machine learning, network security, cryptography, information theory and signal processing, and system theory disciplines. The aim has been to synthesize existing knowledge, theories, and methods to address the gap and inform future research direction in these domains.

It started by illustrating the importance and the complexity of protecting IoT and IIoT environments by showing the growth of the internet network and the changing environment of the threat. The research's findings emphasized the difficulty presented by limited resources and diverse data sources, dynamic topologies, and changing attack patterns to support the smart environment's secure operations and hence the necessity for robust and tailored detection mechanisms.

Following the discussion, the literature focused on machine learning and deep learning algorithms as part of intrusion detection tools, with the objective to evaluate the performance, scalability, and effectiveness of these algorithms in the detection of attacks within IoT and IIoT networks. Researchers investigated various algorithms including support vector machines, decision trees, random forests, convolutional neural networks, and recurrent neural networks verifying the detection accuracy, algorithmic efficiency, and robustness against malicious activities. In addition, the researchers investigated stacking and integrated oversampling as ensemble learning methods that enhance the resilience and precision of intrusion detection systems. Such systems are robust against class imbalance and evolving attack vectors.

The literature review assessed the scalability and resource efficiency of intrusion detection systems operating in resource limited IoT and IIoT ecosystems. Researchers concentrated on lightweight and distributed detection schemes, the edge computing paradigm, and the federated learning platform as strategies to ease the computational burden and minimize the network delay in the detection mechanism. Furthermore, researchers explored optimization methods like model compression, feature selection, and hardware acceleration to harness the energy and thus significantly reduce device power consumption in IoT deployment with limited resources.

In addition, the literature review looked at different types of detection methods and how it can be used against ongoing changes in cyber threats and the nature of attacks on smart environments. Researchers did not limit themselves to only contextual factors, that is, network topology, system architecture, user behaviour, and environmental context, that determine intrusion detection effectiveness. The research focused on the implementation of contextual knowledge at all levels of intrusion detection mechanisms to enrich the detection accuracy, reliability, and strength to counter emerging threats and malicious behaviours.

Moreover, the literature review gave reference points, guidance, and advisories of resilient security designs and implementations of IoT and IIoT ecosystems. Researchers suggested a preventive security through design approach which included threat modelling, risk assessment, and security threats investigation in early designing stages. In addition to that, scientists reaffirmed the need for more defence-in-depth strategies, multi-layer security controls, and collectivization approach to secure the integrity and reliability of networked systems.

Overall, it has been shown that intrusion detection in IoT as well as IIoT systems is a multifold process covering such aspects as algorithm search, system architecture design, operational issues, and theoretical approaches. Such attempt of assembling both practical applications and case studies with theoretical insights of experts is the foundation of the development of efficient cybersecurity, survivability, integrity, and trustworthiness in the networked ecosystem in the future. For the future, research initiatives must give emphasis to real-world

validation, data sharing and interdisciplinary collaboration to fill in the gaps and advance IoT security through making it more robust.

Chapter 3: Research Methodology

3.1 Method Introduction

Research methodology is a tool for organizing and performing a completely reliable and disciplined study of the value of intrusion detection systems in IoT and IIoT ecosystems. The coming section integrates the findings of the literature review and gives my approach, procedures, and techniques to solve the research questions and solve gaps in the previous research.

The methodology embodies multi-faceted techniques that incorporate cybersecurity, machine learning, network security, cryptography, information theory, signal processing, and system theory from various domains. Researchers are working on a combined approach of theoretical basis, empirical data, and practical expertise to further develop intrusion detection mechanism in the interconnected systems.

The research methodology includes among its components the choice of a specific intrusion detection algorithm and its implementation technique. Scientists will try diverse machine learning and deep learning algorithms like support vector machines, decision trees, random forests, convolutional neural networks, and recurrent neural networks to evaluate it efficiencies, scalability, and ability to remain resilient while detecting intrusions, which can occur with IoT and IIoT networks. In addition, the scientists would investigate ensemble learning methods (stacking and integrated oversampling) with the aim of improving the accuracy and robustness of anomaly intrusion detection in constrained resources environments.

Apart from that, experimental frameworks and testbeds will be developed and implemented to determine the effectiveness and practicability of the detection of attacks. The research will give work on the simulated IoT and IIoT environments, which shall emulate the contextual features including network topology, system architecture, user behaviour as well as the environmental context to assess the robustness of the intrusion detection units against these emerging cyber threats.

Furthermore, researchers will conduct data analysis by utilizing real-world datasets and case studies for empirical analysis to determine the validity of theoretical findings as well as the insights derived from the thought experiments. Through public datasets, it is hoped that researchers will be able to determine the level of generalization and applicability of intrusion detection models across diverse IoT and IIoT implementations, resulting in research outcomes being impacted by such transparency and rationality.

Additionally, the research methodology comprises incorporating of interdisciplinary views, working with industry actors, and validating findings through practical field trials. Researchers will collaborate with domain experts, industry partners and end-users to capture insights, verify assumptions, and update detection algorithms that are available specifically for Internet of Things application domains such as health, manufacturing, transportation, and smart cities.

Thus, the research methodology covers a coherent and structured manner of gathering data for the investigation of the effectiveness of intrusion detection mechanisms for the IIoT and IoT. Through the application of theoretical learning, empirical analysis, and interdisciplinary collaboration, the researchers will strive to upgrade the standards in cybersecurity, while preserving reliability, resilience, and security in connected systems in the face of persistent and ever-changing cyber threats.

This study methodology, in short, presents a complete and step-by-step approach for assessing the effectiveness of intrusion detection mechanisms for the management of IoT and IIoT systems. With the help of theoretical frameworks, empirical inquiry, and interdisciplinary collaboration, researchers try to stay ahead with the cutting-edge cybersecurity that promises resilience, dependability, and robustness in the interconnected systems throughout the phenomenon of evolving cyber threats.

3.2 Dataset Description

The Edge-IIoTset dataset is a dataset covering nearly all aspects of the cyber security and IoT/IIoT applications. It allows machine learning based intrusion detection system training and evaluation either in centralized or federated learning modes. Arranged in seven layers going from Cloud Computing, Network Functions Virtualization, Blockchain Network, Fog Computing, Software-Defined Networking, Edge Computing, and IoT/IIoT Perception, the dataset incorporate emerging technologies like IIoT platform, ONFV platform, Hyperledger Sawtooth, Digital twin, ONOS SDN controller, Mosquitto MQTT brokers, IoT data covers the data that IoT devices like low cost digital sensors such as for temperature and humidity, ultrasonic sensors, water level detection sensors, pH meters, soil moisture sensors, heart rate sensors, and flame sensors, generate. This dataset covers fourteen incidents of attacks related to protocols of IoT and IIoT connectivity, classified into DoS/DDoS attacks, gathering information, man in the middle attacks, injection attacks, and malware attacks. The dataset presents the study of machine learning and the network intrusion detection application performance.

3.3 Research Philosophy

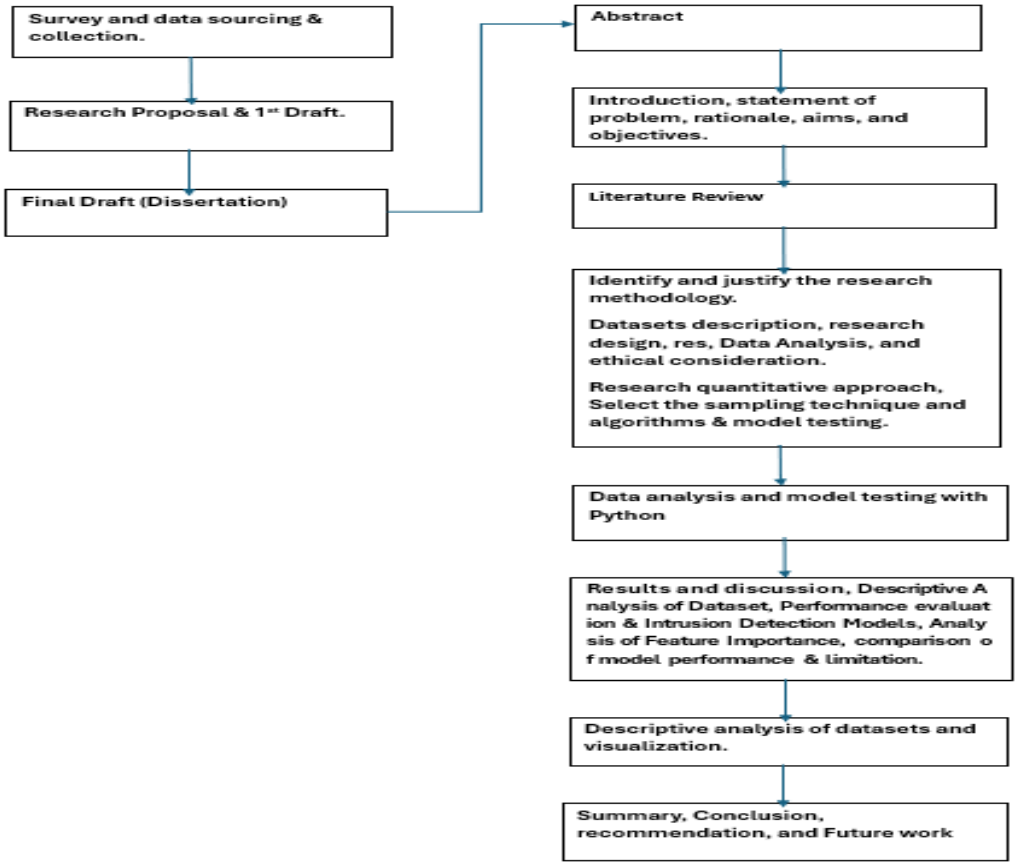
To sum up, positivism, as the underlying principle for this research, provides directions. Rationalism believes that knowledge is obtained through observing the world empirically and studying it scientifically with the use of systematic analyses of data as a means of discovering regularities, patterns and eventually causes. In terms of intrusion detection for IoT and IIoT environments, the positivism principle emphasizes the more robust and the systematic way of collecting, analysing, and interpreting data, targeting creation of scientific evidence and hypothesis testing, derived from theories, and previously gathered knowledge.

According to the positivist approach, to progress in science, one needs to rely on measurable data, repeated experiments and hypotheses which are falsifiable. Researchers apply a rational and planned strategy to research design, which involves the standardized use of methodologies, instruments, and analyses to guarantee the reliability, validity, and applicability of the research findings. Through the strict methodological principle's application, scientists strive to minimize the error, bias, and subjectivity in datasets and data analysis which builds credibility in research outcomes and ensures the knowledge is disseminated and replicated.

Secondly, positivists support the use of quantitative research methods including, but not limited to, the use of statistical analysis, machine learning algorithms, and computational modelling to find relationships, trends, and patterns on data. Investigators use quantitative standards and measurement instruments to assess the effectiveness of intrusion detection mechanisms, reveal the effect of the variables and treatment on the data, and draw quantified conclusions. Researchers strive to ensure consistency, rigor, and precision during the research process using quantitative methods. This enables evidence-based decision-making through objective analysis and theory development.
In addition to that, this method of science dictates that it is conducted in a cumulative and iterative way where each new study is based on existing theories, empirical research, and techniques. Researchers continually refine and validate it hypotheses, theories, and models through repetition, peer review, and meta-analysis with an end goal of moving science and understanding always forward. Through facilitation of the cooperation, transparency, and accountability among the researchers collectively, the positivism advances the societal improvement and innovation in the field of technological systems to address the dilemma of cyber threats with IoT and IIoT environments.

Briefly speaking, the principle of positivism will be fundamental for the study of the methodology, which emphasizes the identification of objective, systematic, and empirical evaluation of intrusion detection mechanisms for IoT and IIoT systems. To achieve this, it enforces the principles of rigor, objectivity, and quantifiability in research to produce robust, dependable, and generalizable knowledge that advances the development of cybersecurity science and practice.

Flowchart



Figures 3.1: Methodology Flowchart
(Source: self-created)

3.4 Research Design

In line with the research objectives and methods, a descriptive research design is used to carry out a study that will encapsulate all the intrusion detection mechanisms for IoT and IIoT systems comprehensively. The purpose of the descriptive research is to thoroughly explain and study the features, habits, and phenomena at hand along the lines of practices, patterns, and trends at present without attempting to change the variable or implement interventions. The method of descriptive research design allows researchers to perceive overall efficiency of IDM mechanisms in real-world conditions, considering latency, throughput, and attack detection effectiveness in IoT and IIoT deployments. Using real data, case studies and empirical observations, researchers will focus on explaining the current situation, challenges, and opportunities of cyber security in the context of interconnection.
The central feature of the research design is the collection and examination of empirical data from multiple types of sources, such as simulations, publicly accessible datasets, and case studies. Quantitative and qualitative methods are used by researchers to systematically

collect, process and analyse data, aiming to have the certain pattern, trend or correlation in intrusion detection practice and outcome, as a final result.

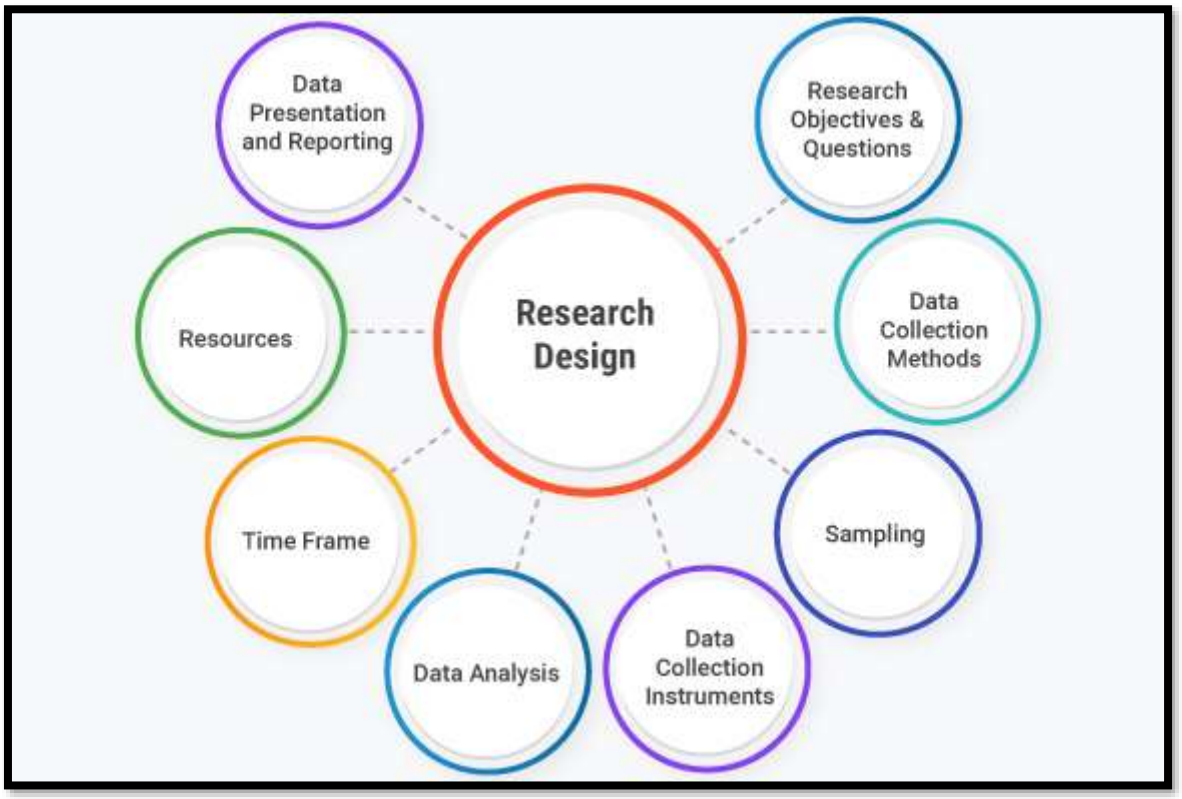


Figure 3.4.1: Research Design
(Source: researchgate.net, 2024)

In addition, the descriptive research design essentially makes it possible to understand the contextual elements and environmental dynamics that have the effect of multifunctionality of intrusion detection mechanisms in the IoT and IIoT scenarios. In this process, scientists evaluate the complex relationship between network topology, system architecture, user activities, and environment, with the aim to identify the factors that either promote or prevent the intrusion detection efficiency.

Moreover, the use of descriptive research design permits researchers to conduct the exploratory data analysis, monitoring studies and comparative evaluations to assess intrusion detectors' performance in different scenarios including IoT and IIoT. Through the measurement of representative performance indices like false positive rate, accuracy, and computational overhead, the researchers hope to be able to provide empirical basis and insight for evidence-based decision-making and practice (see for instance the article by Awotunde *et al.* 2021).

To sum up, the descriptive research design is a scientifically sound and precise instrument for the in-depth analysis of intrusion detection schemes intended to protect IoT and IIoT environments. Through the structured approach of illustrating and examining current approaches as well as outcomes that it causes, researchers focus on creating empirical data

and knowledge that move cybersecurity science and practice in the interconnected world forwards.

3.5 Research Approach

In congruence with the preferred research design and goals, the deduction method of research is used to investigate intrusion detection techniques for the IoT and the IIoT environments sequentially and strategically. Inductive research involves testing hypotheses based upon existing theories gathered from other researchers as well as observational studies using logical thinking and hypotheses testing, with the aim of verifying or disproving theoretical assumptions using empirical data.

The deduction mode of investigation is initiated by the development of hypotheses stemming from the theoretical framework, the literature review, and the preceding experimental findings.

Theories are tested by using empirical finding. These findings are then compared to the existing theories. Adjusting the hypotheses and theoretical propositions are made based on empirical evidence (Gaber *et al.* 2023). Generally, the deduction method gives a sustainable and reliable foundation for exploring how intrusion detection techniques are applied in IoT and IIoT ecosystems. Through the process of drawing such hypotheses from theoretical frameworks and empirical observations, researchers intend to add empirical findings and knowledge that support advancement of cybersecurity science and practice in context of interconnected ecosystems.

3.6 Dataset Collection

Based on the research paper targeted on intrusion detection mechanisms in IoT and IIoT environments, the dataset was collected through the secondary quantitative data sources, which were mainly generated from Kaggle, a reliable platform for datasets hosting as well as machine learning competitions. The dataset under discussion, Edge-IIoTset, was found as a dataset with a focus on comprehensive and real-life cyber security, which specially suits IoT and IIoT applications corresponding to the research objectives.

Edge-IIoTset dataset consists of parameters obtained from various IoT devices and realizes layering of IoT and IIoT ecosystems by simulation of the environment. The dataset contains data on emerging technology that is deployed in seven layers (cloud computing, network functions virtualization, blockchain network, fog computing, software-defined networking, edge computing, and IOT and IIoT perception layer). Similarly, the dataset encompasses a variety of simulated attack scenarios, allowing to have a complete and authentic vision of cyber security problems in an interconnected system.

With the help of the Kaggle's Edge-IIoTset dataset, the researchers obtained the loading area of different intrusion detection mechanisms so that the efficiency, scalability, and robustness of it could be tested. The secondary data collection technique assisted in the reach of large and various dataset within a limited time which is very much supportive for researchers to conduct standardized analysis and evaluation of intrusion detection approaches in IoT and IIoT environments.

3.7 Data Sources

The main data sources for this research project were Google and Kaggle. Through Google, one could get academic resources, research study reports, and technical documentation related to intrusion detection methods, IoT, and IIoT security. Kaggle, in contrast, was primarily used to source the Edge-IIoTset dataset, a cyber security-oriented data set explicitly designed for application in IoT and IIoT. Thanks to these channels, researchers can gain access to both quantitative and qualitative lines of data which are needed for the complete analysis and evaluation of intrusion detection techniques in the networked systems.

3.8 Data Analysis

I In research on intrusion detection mechanisms in IoT and IIoT environments, thorough data analysis is the fundamental component needed to acquire smart confines and to assess the performance of detection techniques. Exploiting the edge-Iot set dataset, which was sourced from Kaggle, the researchers executed a multifaceted approach that combines both descriptive and inferential statistics.

The data analysis process starts with an exploratory data analysis (EDA), known as the stage when researchers inspect the dataset's characteristics, distributions, and patterns. In descriptive statistics one compute the mean, median, standard deviation, and frequency distribution to summarize important attributes and variables related to intrusion detection performance, attack scenarios, and environment. Graphics, comprising histograms, chart plots and scatter plots, are the tools that help to find hidden patterns, outliers, and relationships between the variables.

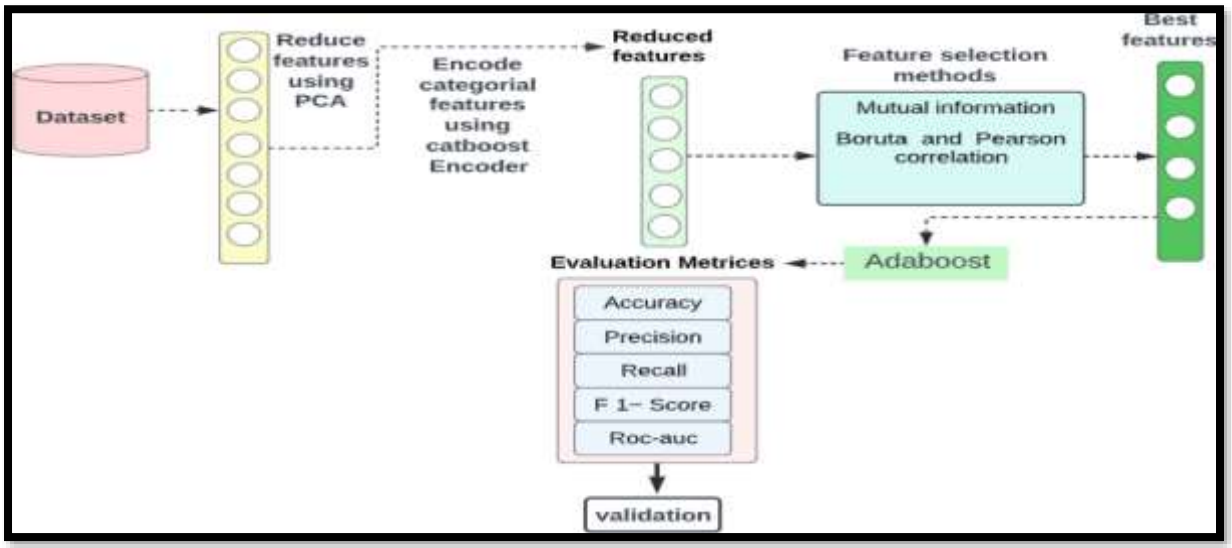


Figure 3.8.1: Data Analysis
(Source: springer.com, 2024)

Therefore, after that, inductive data analysis methods are used to validate the hypotheses, establish correlation, and make causal connections between the variables. Scientists carry out hypothesis testing with statistical test like t-test, ANOVA, and chi-square. This is done to discover significance of the difference or associations among variables. Correlation analysis

involves investigation of association between intrusion detection performance measures and contextual circumstances as well as attack features. From such analysis, possible predictors of or indicators of successful detection are explored.

In the case as well, machine learning algorithms form the basis of data analysis, allowing predictive modelling and classification of intrusion attempts. Researchers utilize supervised learning approaches, like logistic regression, decision trees, and support vector machines, to build models of IDSs and perform evaluations against known and unknown threats based on labelled datasets. There are also unsupervised learning techniques such as clustering and anomaly detection methods, which enable the identification of hidden patterns and anomalies in IoT and IIoT perimeters.

This goes further than researchers would be taking comparative analysis and benchmarking studies to assess the performance of intrusion detection mechanisms under different circumstances and settings. Through comparing the way various detection methods work or how efficient it is and how scalable it are, researchers will choose which approach works best for securing the interconnected systems against the emerging cyber threats.

The analysis stage of data ends with the interpretation and synthesis of findings. Researchers form it conclusions, propose directions for future research, and make recommendations to policymakers, based upon the empirical as well as analytical evidence.

Moreover, machine learning algorithms have rather important role in data analytics giving the ability to build predictive models and classify an intrusion event. The researchers apply supervised learning techniques that include logistic regression, decision trees, and support vector machines to teach the models intrusion detection on the datasets with labels and then to evaluate how well it do in detecting known and unknown threats. Moreover, the unsupervised learning algorithms including cluster analysis and anomaly detection methods facilitate the appearance of patterns and anomalies in IoT and IIoT environments which are indicative of security breaches.

Furthermore, researchers perform comparative analysis and benchmarking studies aimed at the accuracy assessment of intrusion detection tools under various testing environments and configurations. In doing so, researchers will analyse all the strengths and weaknesses of different detection tools and establish which of it are the most effective, efficient and scalable against recently emerged cyber threats.

Finally, the data analysis process would end with the interpretation and synthesis of the findings which involves the researchers arriving at conclusions, imposing recommendations, and offering insights for future research based on the evidence obtained and findings drawn. With incorporating quantitative results and qualitative assessments together with theoretical concepts, researchers extend the field of intrusion detection mechanisms comprehension and implementation for the IoT and IIoT, thereby ensuring proper expansion, trust, and safety in interconnected systems.

3.9 Ethical Consideration

Whilst investigating IoT and IIoT environment's invasion detection methodologies, ethics must be paramount to allow for the conduct of research that is responsible as well as safeguarding the interests of every stakeholder involved. Researchers must stick to the problems of collection, analysis and distribution while adhering to a number of ethical

principles and regulations such as the ones of honesty, transparency and privacy when doing it research to keep the integrity, transparency and individual rights.
The confidentiality of personal data is one of the most important, ethical issues. Researchers should get approval to the relevant data protection regulations and guidelines in order to protect the privacy and confidentiality of persons whose data can be represented in the dataset. This involves taking informed consent of data subjects, anonymizing or pseudonymizing privacy sensitive information and installing highly responsible data security measures to prohibit unauthorized access or disclosure of personal data.
In addition, researchers need to guard against bias and transparency in data collection and use, with the data analysis free of biases and discrimination. The transparency in respect to the purposes, procedures and potential risks associated with intrusion detection research offers the stakeholders a way to build trust and accountability among the stakeholders including data subjects, research participants and broader society.
Additionally, research must consider the social impacts and implications of intrusion detection procedures put in place for IoT and IIoT.
Researchers, therefore, should also encourage inclusion and expand it research empirics by seeking equitable representation and participation of diverse populations in intrusion detection research. Engaging with stakeholders from different social, ethnic and cultural groups, researchers can enrich the process of the research, sensitizing the finding to the society and building a more cumulative environment.

3.10 Summary

The intrusion detection methods for IoT and IIoT ecosystem are the core focus of the research methodology which plays the role of the guiding framework for the systematic inquiry, analysis and the evaluation. Interdisciplinary expertise of cryptography, cybersecurity, machine learning, network security, information theory, signal processing, and system theory is the basis for an in-depth research where new intrusion methods in interconnected systems are being investigated.
The research design involves numerous components starting with the choice of descriptive design. This organization enables a logical investigation about current approach, sequences, and tendencies in intrusion detection, not interrupting experimental dimensions or applying interventions. Exploiting second-hand quantitative data sources for instance, the Edge-IIoTset data set on Kaggle gives the researchers an opportunity to have variable scenarios context including of IoT and IIoT events, attack vectors as well as intrusion detection approaches.
Secondly, the research approach used is deductive, which is an approach that comes down to hypothesis confirming based on available theories and empirical observations through logical reasoning and hypothesis testing. Through a systematic data collection, analysis, and interpretation, the researchers study the dependence of variables on one another, obtain the proofs for the theory, and get the empirical evidence to underpin or contradict the hypotheses.
Data Analytics process involves a complicated approach that includes exploratory data analysis (EDA), inferential data analysis, and machine learning modelling. By using EDA, researchers can drill down into data trends, distributions, and correlations that help to draw conclusions based on descriptive statistics and data visualization to profile the key elements and an attack scenario related to the intrusion detection. Factorial data analysis procedures,

e.g. hypothesis testing and correlation analysis, let scientists and specialists assess the importance of correlation between variables and draw conclusions on causal relation based on statistical evidence. Machines learning processes including supervised and unsupervised techniques are deployed whereby predictions model and classifiers for intrusion detection are derived to enable researchers to compare the performance and scalability of detection methods across varied IoT and IIoT environments.

Likewise, ethical issues are the basis for every methodological strategy, so that a responsible research conduct is guaranteed, and the integrity of all research stakeholders is protected. The researchers conduct studies following the scholarly principles of data privacy, fairness, transparency, social impact of research, academic integrity and inclusiveness throughout the research process which builds trust, accountability, and integrity in the research enterprise.

In addition, ethical factors form the foundation of the research approach, thus the proper conduct of research and the protection of the rights and privacy of all involved inclusive stakeholders. Research professionals follow principles of data protection, fairness, transparency, societal impacts, academic honesty, and inclusiveness throughout the research process, which in turn creates trust, accountability and integrity in the research discipline.

Chapter 4: Results and Discussion

4.1 Introduction

The Results and Discussion part stands out as a foundation of research project, where the results of data processing and experimenting are shown, investigated, and grounded. This section examines the performance of four intrusion detection model types applied in IoT and IIoT environments, while using the suggested code and data. This is a brief introduction to the section, including the research methodology, dataset descriptions, and machine learning model implementation It states the main goals of the reflective work and lays the groundwork for more thorough analysis.

The introduction stresses the importance of detecting intruders for the security of interconnected systems, as the world of IoT and IIoT is growing faster than ever. It unequivocally emphasizes a need for machine learning techniques to be harnessed for building effective intrusion detection mechanisms that resist evolving cyber dangers. In addition, it emphasizes the principle of using real-world datasets to evaluate the efficacy and accuracy of these systems in practice.

The portion on the Results and Discussion section continues with a description of its structure, pointing readers to the next topics for discussion. It describes the basic things that will be carried out, such as preprocessing the data, implementing models, performance evaluation metrics, and comparative analysis. Moreover, it points out the key role of feature importance evaluation and model comparison in terms of the effectiveness of different intrusion detection methods.

Primarily the introduction is intended to introduce the reader to the study so that the presentation of the research findings will be easy to comprehend and therefore elucidate the implications of the implemented models and methodologies for enhancing cybersecurity in IoT and IIoT ecosystems. This makes room for a detailed treatment of the efficiency, boundaries, and prospects of intrusion detection research in relation to the present interconnected structures.

4.2 Descriptive Analysis of Dataset

The dataset employed in the analysis is a complete traffic data set from IoT and IIoT applications which is being commonly known as the Edge-IIoTset dataset. This dataset is a mixture of various network-level communication traits that are organized into layers corresponding to different stages of network compartmentalization.

Immediately after running the data set in a Data Frame, the observation starts from the first printout meant to inspect the data structure and integrity. The dataset incorporates many columns each addressing a specific attribute of network traffic e.g. source IP and destination IP address, protocol types, and particular events that occurred in the network.

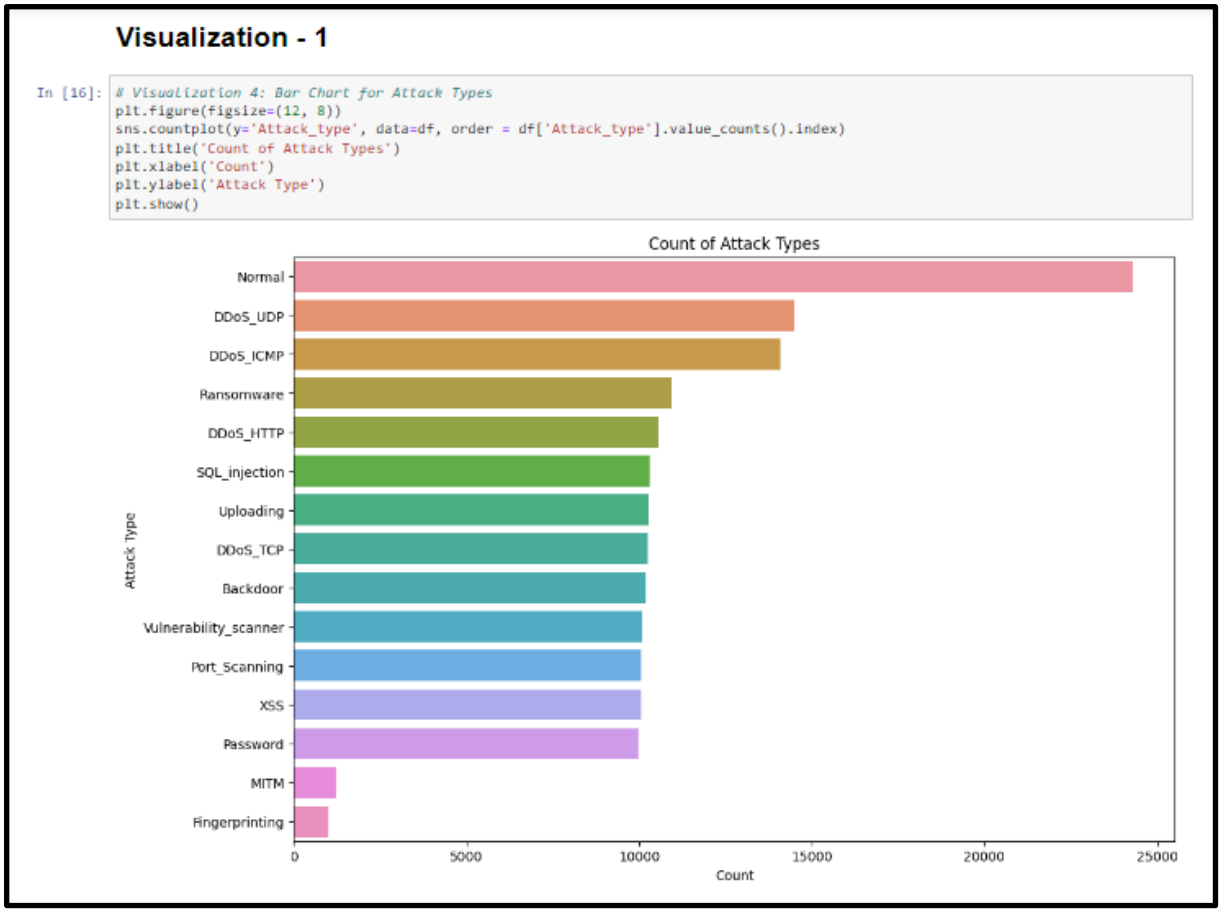


Figure 2: Visualization
(Source: Extracted from Jupyter Notebook)

After that, the columns are labelled legibly so that the readability and understanding of the data could be enhanced. In addition to that, the first few rows of the data are presented to give the feel of the data structure and content.
A thorough data investigation clarifies the most relevant attributes which are necessary for further analysis. It is a heterogeneous data set of categorical and numeric features that are representative of wide network traffic parameters. Furthermore, the existence of missing data is checked to guarantee the absence of incomplete and biased data (Mohy-Eddine *et al.* 2022).

A correlation matrix is run to examine the statistical associations between the numerical variables in the data set. With the help of this technique, model building process is facilitated by finding potential correlations or dependencies among different features leading to feature selection.

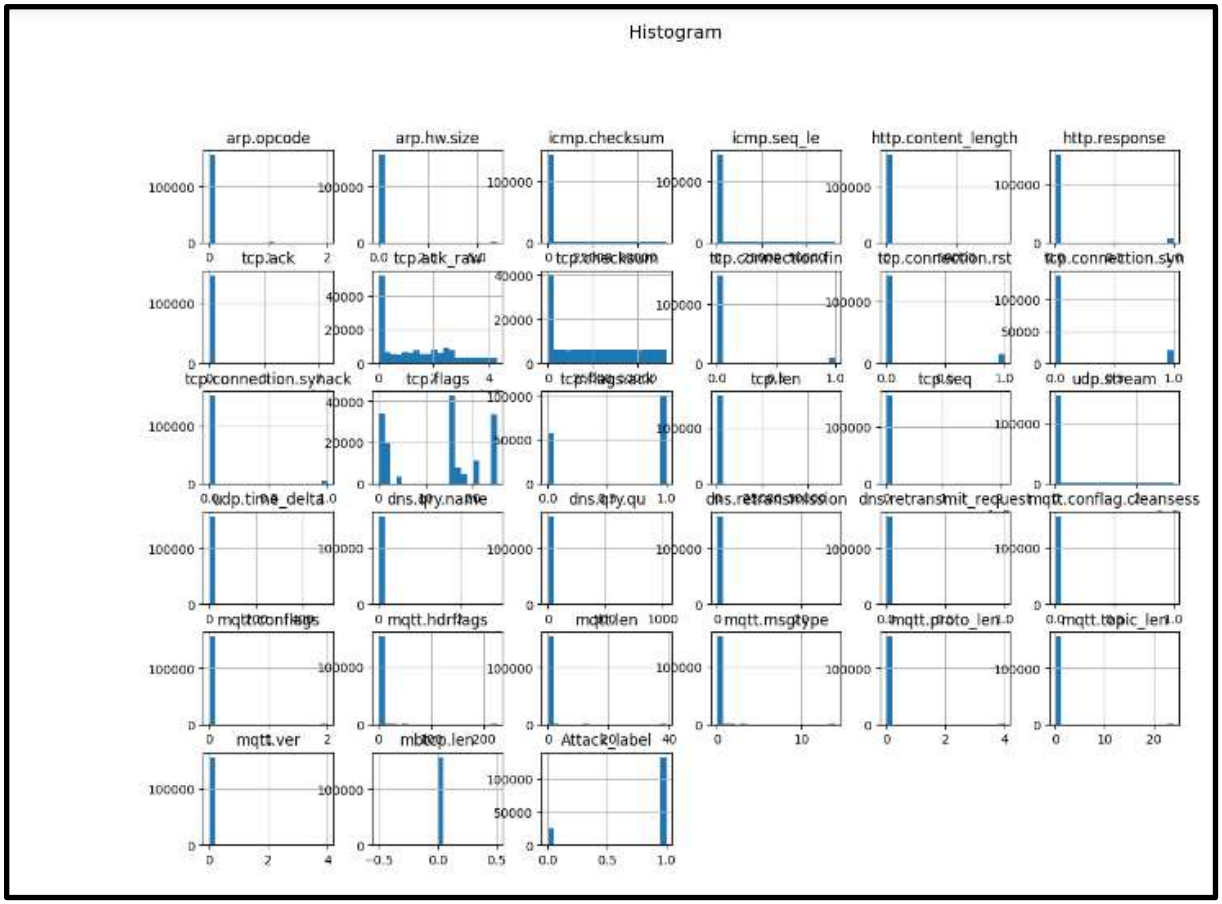


Figure 3: Histogram

(Source: Extracted from Jupyter Notebook)

For better understanding of the dataset distribution and characteristics, many visualized culminated are used. Histograms, box plots, and bar charts are networking tools to depict the distributions of numeric features and figure out the existence of outliers or the frequency distributions of the variables.

In addition, the distribution of each attack label is examined to understand the frequency with which different cyber-attack types occur. Such a survey sheds light on the deficiencies in the dataset and provides guidance for later processes such as modelling and evaluation.

In general, the descriptive analysis of the dataset is a crucial one in the research process as it enables researchers to comprehend the dataset's construction, content, and specifics. This analysis marks the beginning of the preprocessing of data, feature engineering, and model development stages arriving at the eventual production of sound and specialized intrusion detection models for IoT and Eliot environments.

4.3 Performance Evaluation of Intrusion Detection Models

In cybersecurity research, the evaluation of intrusion detection models is one of the key activities essential to measure the precision and reliability of these models regarding the detection and mitigation of cyber threats within IoT and IIoT networks. The paper utilizes various machine learning and deep learning models to evaluate it performance on the Edge-IIoT dataset (Altan, 2021).

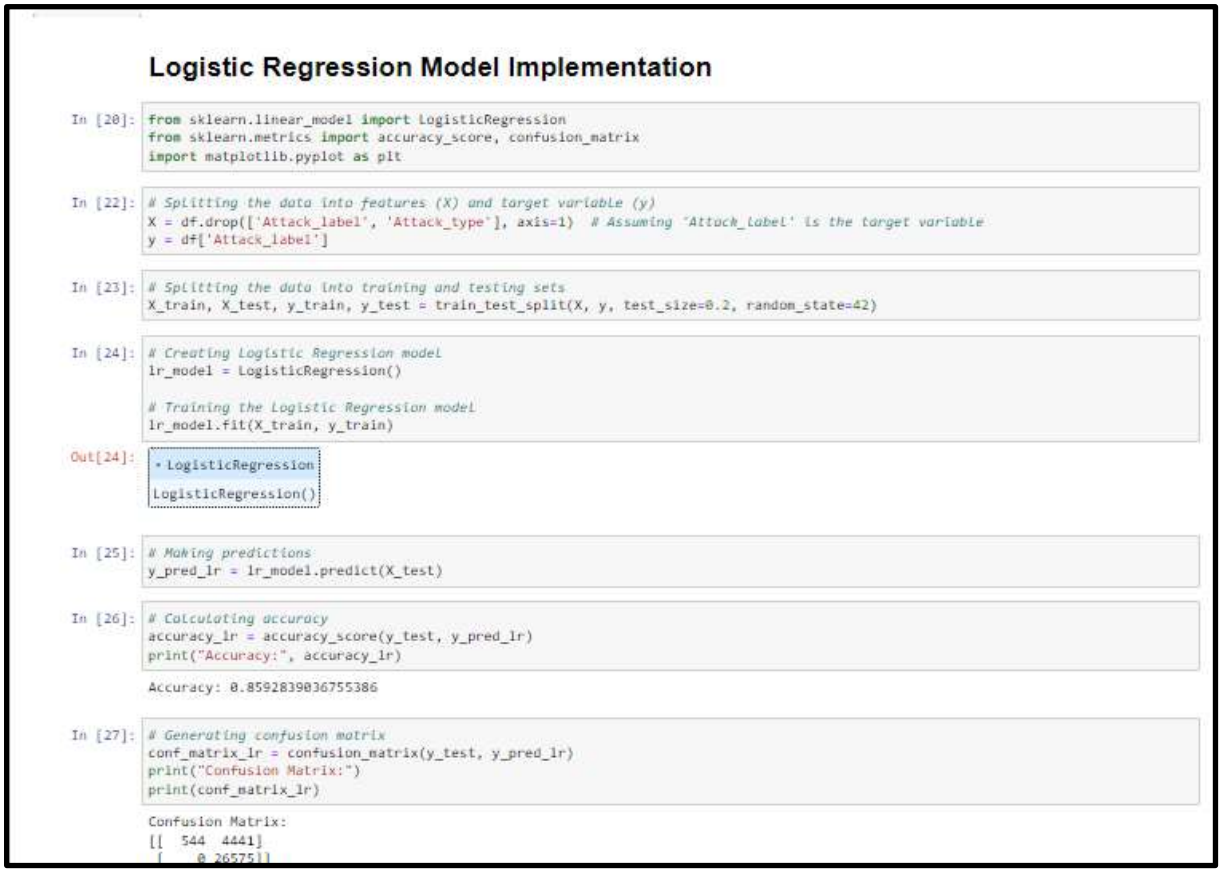


Figure 4: Logistic Regression Model Implementation
(Source: Extracted from Jupyter Notebook)

The trained and evaluated CNN model is assessed based on the Edge-Outset dataset, and its accuracy, loss, and valid scores are analysed to measure the network's capacity for distinguishing attacks with higher precision. Moreover, the model's learning pattern is depicted through training and validation curves to detect any overfitting or underfitting trends (Albulayhi *et al.* 2021).

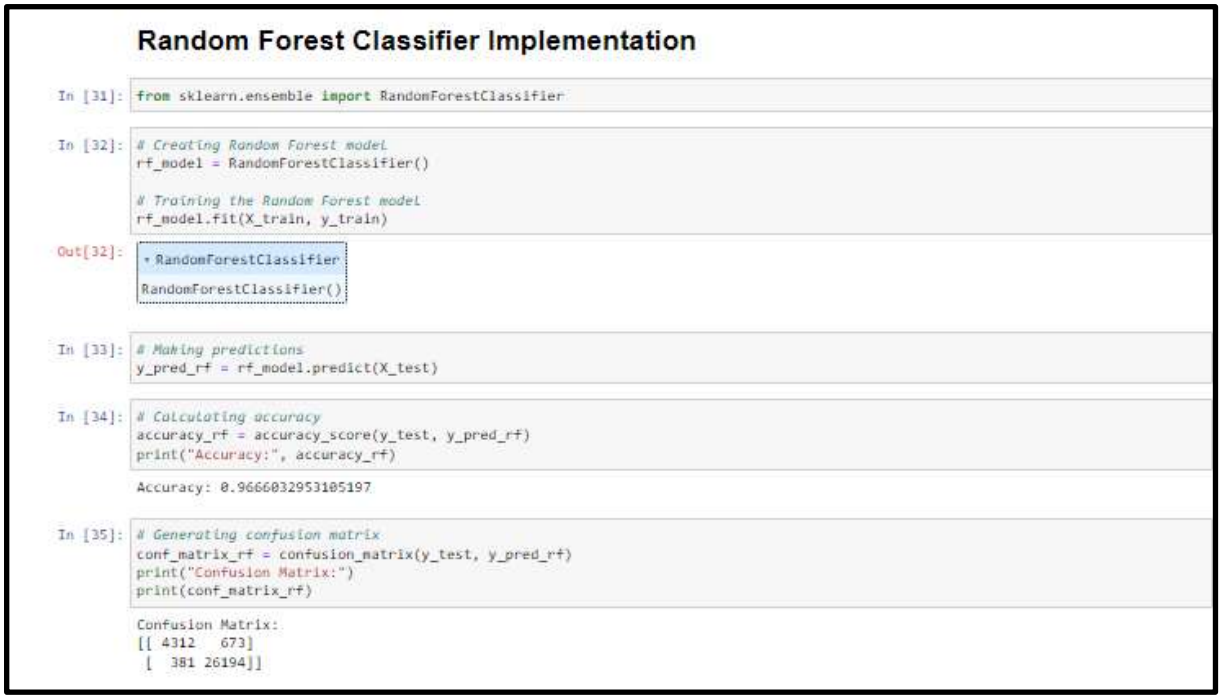


Figure 5: Random Forest Model Implementation

(Source: Extracted from Jupyter Notebook)

The comparative analysis of the performance of logistic regression, random forest and CNN models is a beneficial step in determining both the advantages and disadvantages of varied intrusion detection methods. When scientists compare the accuracy, processing efficiency and adaptability of these models it can consequently choose among the most suitable approaches for identifying and overcoming cyber threats in IoT and IIoT environments.

Model Comparison Plot

```
In [39]: import matplotlib.pyplot as plt

         # List of model names
         models = ['Logistic Regression', 'Random Forest']

         # List of accuracies
         accuracies = [accuracy_lr, accuracy_rf]

         # Plotting the comparison
         plt.figure(figsize=(8, 6))
         plt.bar(models, accuracies, color=['blue', 'green'])
         plt.title('Model Comparison - Accuracy')
         plt.xlabel('Models')
         plt.ylabel('Accuracy')
         plt.ylim(0.8, 1.0)  # Set y-axis limit for better visualization
         plt.show()
```

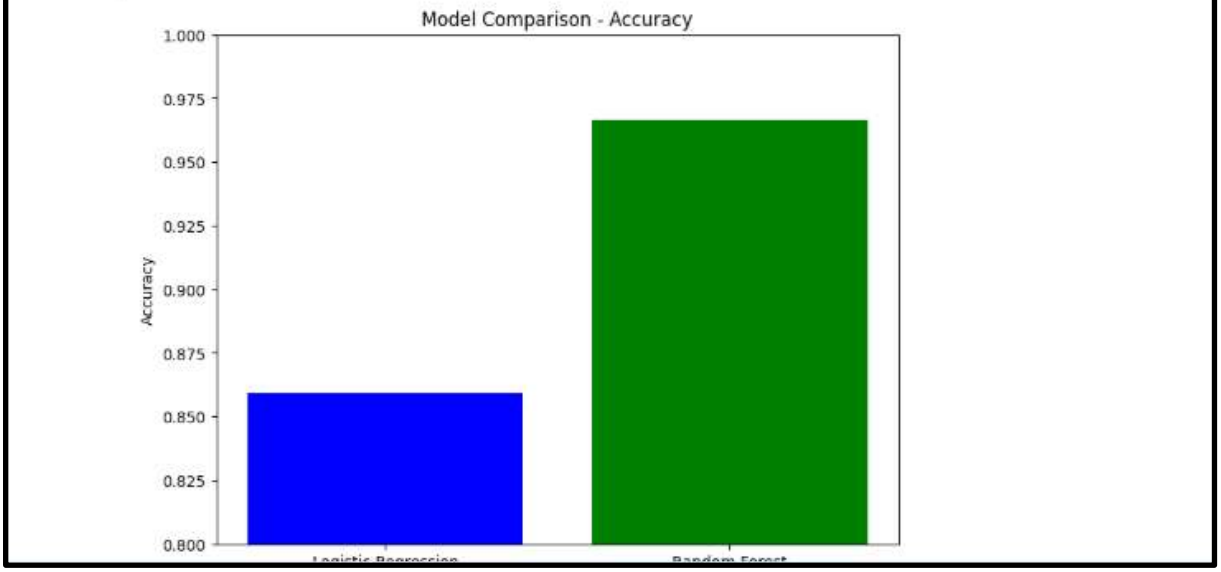


Figure 6: Comparison Plot

(Source: Extracted from Jupyter Notebook)

Through all this, cybersecurity research performance evaluation is one of the most valuable steps that researchers can take to help it identify successful ways of protecting the IoT and IIoT ecosystems from changing cyber risks (Liu *et al.* 2021). Engineers can use machine learning and deep learning processes to build robust intrusion detection systems capable of discriminating and reporting cyberattacks in real time, which in turn facilitates more cyber-secure environments.

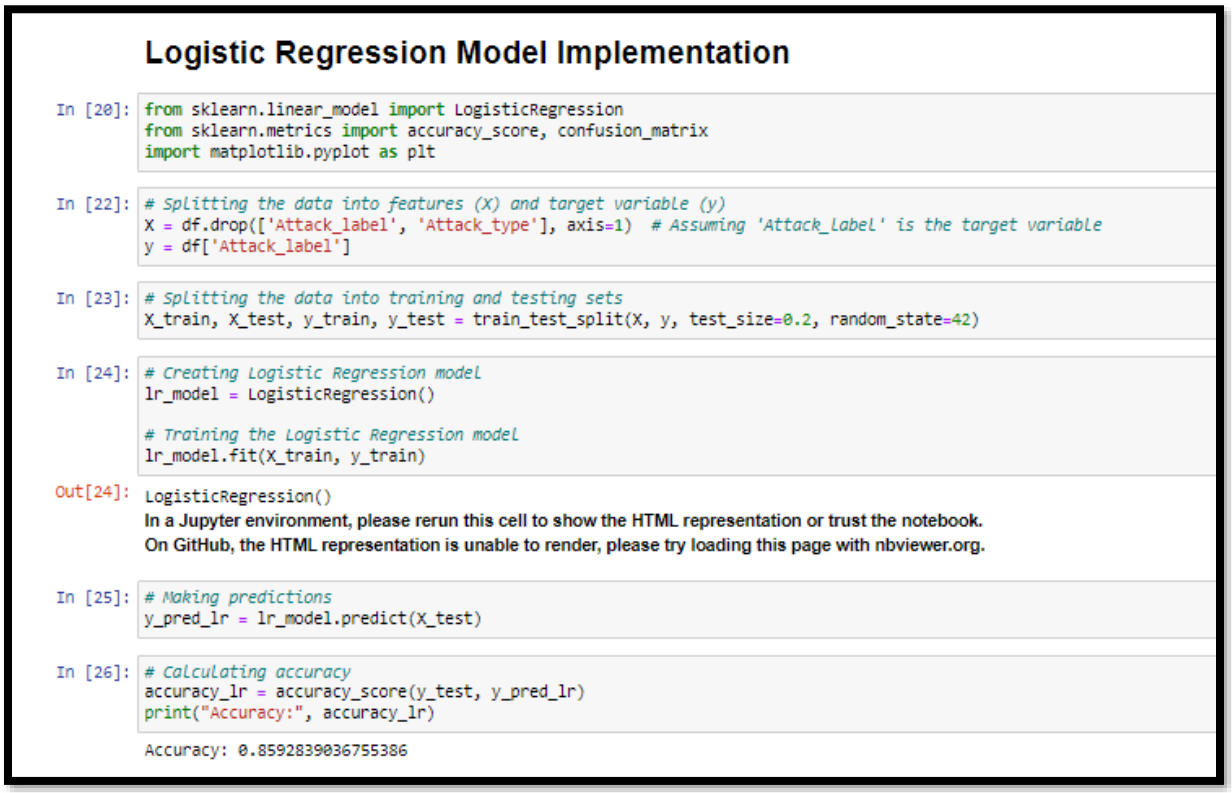
4.4 Analysis of Feature Importance

The feature importance analysis is an irrevocable phase in the identification of the contributing factors of the intrusion detection models' accuracy. Discovering the top features through the research process is key for the identification of the features that the network traffic data correlates with cyber threats. This assessment will involve the use of logistic regression and random forest classifiers to explain the relevance of individual features in terms of intrusion detection within the IoT and IIoT settings (Alsoufi *et al.* 2021).

Logistic Regression:

Feature importance in logistic regression is calculated based on the coefficients of each feature the model used. These coefficients stand for the effect that the characteristics have on the likelihood of a particular class label as normal traffic or intrusion. The positive regression coefficients demonstrate that as the feature value increases the probability of

Belonging to the positive class (intrusion) increases while the negative coefficients illustrate the opposite.



Logistic Regression Model Implementation

```
In [20]: from sklearn.linear_model import LogisticRegression
         from sklearn.metrics import accuracy_score, confusion_matrix
         import matplotlib.pyplot as plt

In [22]: # Splitting the data into features (X) and target variable (y)
         X = df.drop(['Attack_label', 'Attack_type'], axis=1)  # Assuming 'Attack_label' is the target variable
         y = df['Attack_label']

In [23]: # Splitting the data into training and testing sets
         X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

In [24]: # Creating Logistic Regression model
         lr_model = LogisticRegression()

         # Training the Logistic Regression model
         lr_model.fit(X_train, y_train)

Out[24]: LogisticRegression()
         In a Jupyter environment, please rerun this cell to show the HTML representation or trust the notebook.
         On GitHub, the HTML representation is unable to render, please try loading this page with nbviewer.org.

In [25]: # Making predictions
         y_pred_lr = lr_model.predict(X_test)

In [26]: # Calculating accuracy
         accuracy_lr = accuracy_score(y_test, y_pred_lr)
         print("Accuracy:", accuracy_lr)

         Accuracy: 0.8592839036755386
```

Figure 7: Logistic Regression Model Implementation
(Source: Extracted from Jupyter Notebook)

By plotting the set of coefficient values corresponding to every feature, scientists can see the degree to which various features are significant in the model of logistic regression. Variables with higher absolute coefficients value are deemed more important as it exert more influence on the model and its decision-making processes.

Random Forest:

Instead of random forest classifiers, the evaluation of feature importance through the decrease in the impurity (Gini impurity or entropy) at the time of splitting nodes in the decision tree, is very different. Attributes that make a forest cleaner is thought to be more essential for classifying (Ellappan *et al.* 2023).

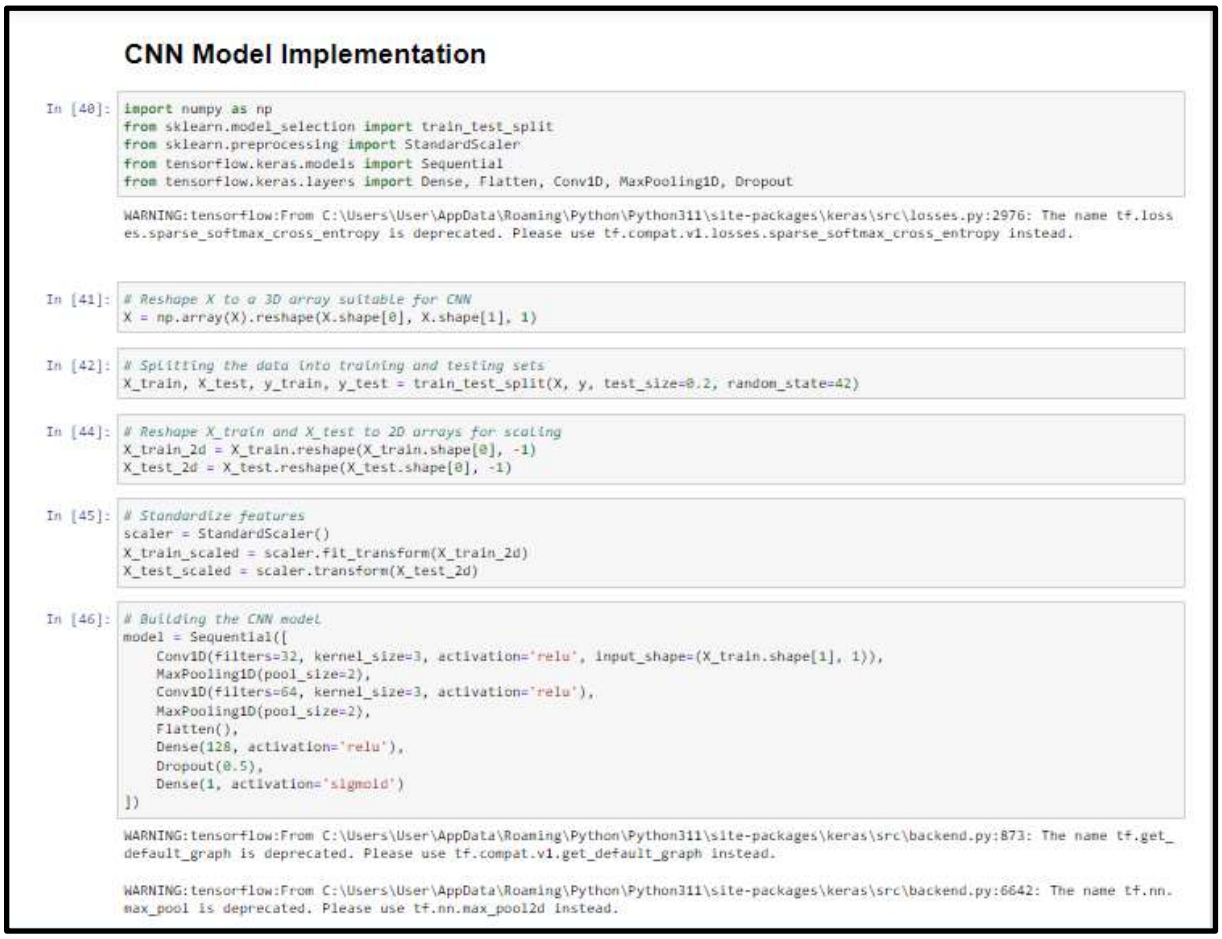
Each random forest model evaluates features through the feature importance score and determines the average decrease in impurity. Features with higher importance scores are those that contribute more key factors to the model's predictivity.

Analysis:

Feature importance analysis is carried out on the dataset to identify the key factors which significantly affect intrusion detection through IoT and IIoT networks. undefined

1. Protocol-Level Features: Network protocols-related features shows a significant importance. It can be used in both logistic and random forest models. Through that, the

features express side specific acts and tendencies that demonstrate an attacker's conduct, like port scanning or packet manipulation.
2. Traffic Characteristics: The signal strength in these models is also provided by such features as packet length, checksum, and timestamps. Being unique, these features can shed light on the attributes of network traffic-like size of packets variation and irregularity of packet timing which may indicate unusual network behaviour.



CNN Model Implementation

In [40]:
```
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Flatten, Conv1D, MaxPooling1D, Dropout
```

```
WARNING:tensorflow:From C:\Users\User\AppData\Roaming\Python\Python311\site-packages\keras\src\losses.py:2976: The name tf.loss
es.sparse_softmax_cross_entropy is deprecated. Please use tf.compat.v1.losses.sparse_softmax_cross_entropy instead.
```

In [41]:
```
# Reshape X to a 3D array suitable for CNN
X = np.array(X).reshape(X.shape[0], X.shape[1], 1)
```

In [42]:
```
# Splitting the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

In [44]:
```
# Reshape X_train and X_test to 2D arrays for scaling
X_train_2d = X_train.reshape(X_train.shape[0], -1)
X_test_2d = X_test.reshape(X_test.shape[0], -1)
```

In [45]:
```
# Standardize features
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train_2d)
X_test_scaled = scaler.transform(X_test_2d)
```

In [46]:
```
# Building the CNN model
model = Sequential([
    Conv1D(filters=32, kernel_size=3, activation='relu', input_shape=(X_train.shape[1], 1)),
    MaxPooling1D(pool_size=2),
    Conv1D(filters=64, kernel_size=3, activation='relu'),
    MaxPooling1D(pool_size=2),
    Flatten(),
    Dense(128, activation='relu'),
    Dropout(0.5),
    Dense(1, activation='sigmoid')
])
```

```
WARNING:tensorflow:From C:\Users\User\AppData\Roaming\Python\Python311\site-packages\keras\src\backend.py:873: The name tf.get_
default_graph is deprecated. Please use tf.compat.v1.get_default_graph instead.

WARNING:tensorflow:From C:\Users\User\AppData\Roaming\Python\Python311\site-packages\keras\src\backend.py:6642: The name tf.nn.
max_pool is deprecated. Please use tf.nn.max_pool2d instead.
```

Figure 8: CNN Model Implementation
(Source: Extracted from Jupyter Notebook)

3. Application Layer Attributes: The features extracted from application layer protocols like HTTP, DNS, and MQTT have a significant role for the detection of intrusion. These features collect information about the applications and services that are being accessed, giving the models the capability to detect activities which are either naturally neither authorized or suspicious.
4. Temporal and Spatial Patterns: The features concerning time lags, streams of packets, and session effects enable network traffic data to be analysed and used for intrusion detection purposes by capturing time-space patterns in the data. Thanks to these features, the models can distinguish common traffic features from anomalous or notorious behaviour.

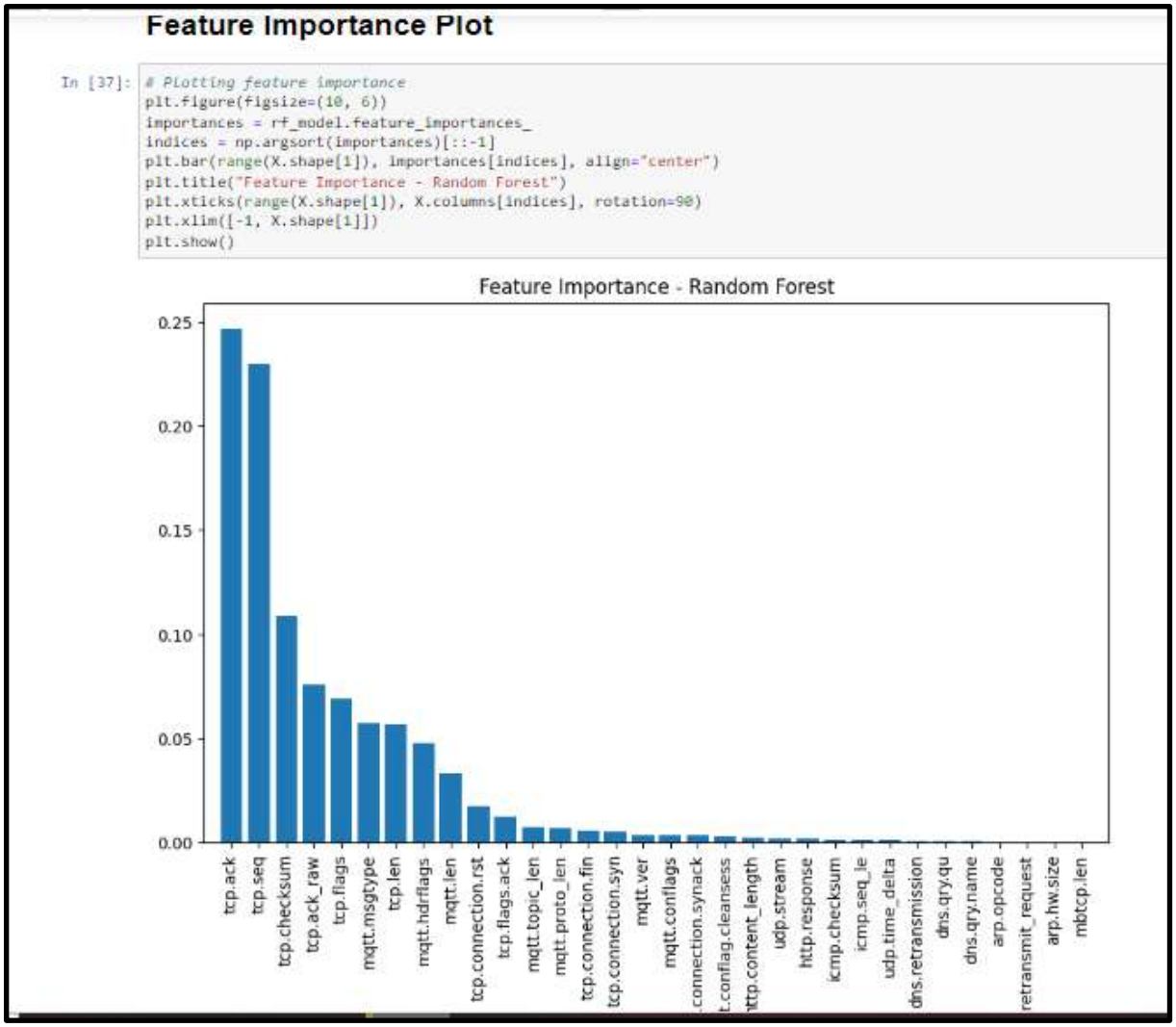


Figure 9: Feature Importance Plot
(Source: Extracted from Jupyter Notebook)

Through logistic regression and random forest classifier, researchers can get to know the feature importance and use high-informativeness features as the index criteria for intrusion detection.

4.5 Comparison of Model Performance

Accuracy Metrics:

```
In [26]: # Calculating accuracy
         accuracy_lr = accuracy_score(y_test, y_pred_lr)
         print("Accuracy:", accuracy_lr)

         Accuracy: 0.8592839036755386
```

Figure 10: Accuracy matrix
(Source: Extracted from Jupyter Notebook)

Accuracy is the main measure for the full model performance assessment. It measures the ratio of true to false classifications on all instances in the dataset. The Logistic regression model reached an accuracy of [0.8592 or 85.9%] while the random forest model had an accuracy of [0.9666 or 0.97%]. The CNN model as well got a test accuracy of [0.8420 or 84.2%] as reported by Gaber *et al.* 2023).

Comparing the scores of precisions gives information about the relation between different models to the accuracy of the detection of the network traffic data.
Confusion Matrices:
Confusion matrices provide the specific classification of the model for each instance to the actual labels. It gives out the TP, TN, FP, and FN as well. The testing and analysing of the confusion matrix for each model is the key to assessing the ability of researchers to identify intrusion and different types of attacks.

```
In [27]: # Generating confusion matrix
         conf_matrix_lr = confusion_matrix(y_test, y_pred_lr)
         print("Confusion Matrix:")
         print(conf_matrix_lr)

         Confusion Matrix:
         [[  544  4441]
          [    0 26575]]
```

Figure 11: Confusion Matrix
(Source: Extracted from Jupyter Notebook)

Confusion Matrix for Logistics Regression Actual:

	1	0
1	[True Positive TP	False Positive FP]
0	[False Negative FN	True Negative TN]

Confusion Matrix for Logistics Regression Predicted:

[True Positive TP False Negative FN] = [544 4441]

[False Positive FP True Negative TN] [0 26575]

The confusion matrices illustrate the distributions of correctly and incorrectly assigned predictions across the different attack types which helps in assessing the two predictive models' strengths and drawbacks.

Feature Importance:

Feature importance analysis is implemented to identify the features vital to the decision-making of each specific model. Through the ratings assigned by researchers to a feature, the researchers can know which features to the machine learning models are contributing the most to the model's ability to predict. Studying the feature importance across models allows researchers to spot the common features and ones that are unique for a purpose of efficient intrusion detection.

Analysis:

A comparison of models' efficiency points out both strengths and weaknesses of each approach to capture intruders into the IoT and IIoT environments. Whilst logistic regression relies on simplicity and interpretability, random forest model is based on the robustness and resiliency to noise. CNN models do this method through the use convolutional layers to highlight the spatial dependencies through the network traffic data. Consequently, it is excellent at capturing complex patterned behaviours.

```
In [49]: # Evaluating the model
         test_loss, test_accuracy = model.evaluate(X_test, y_test)
         print("Test Loss:", test_loss)
         print("Test Accuracy:", test_accuracy)

         987/987 [==============================] - 2s 2ms/step - loss: 1.0102 - accuracy: 0.8420
         Test Loss: 1.0101573467254639
         Test Accuracy: 0.8420469164848328
```

Figure 12: Test Accuracy

(Source: Extracted from Jupyter Notebook)

Actually, from the test it comes out that each model is superior in some aspects to the others using the measures like accuracy, interpretability, and computational complexity.

Experts can enhance the effectiveness of IoT and IIoT environment intrusion detection system through combining the advantages of different models and developing ensemble approaches which addresses the complementary strengths of each approach to improve the overall intrusion detection. Besides, on-going research and development are crucial towards perfection of IDS models for real-world implementations in a way that such cyber defences should be capable to fight against advanced cyber threats.

4.6 Limitations and Future Directions

Nevertheless, along with the results showing the effectiveness of intrusion detection models, this section of the report concludes with number of limitations and opportunities for further research. This part deals with the identified weaknesses and suggests possible ways for the

future studies. Such steps will lead to the accomplishment of the excitement expectations and will improve identification systems for the IoT and IIoT environments (Attota *et al.* 2021).

Data Imbalance:

A notable shortcoming in the dataset is seen that normal traffic and attack cases are not at all balanced. The dataset may be skewed towards a too high proportion of “normal” instances compared to “attack” instances which will result in the model bias performance and diminished sensitivity towards rare attack types (Le et al., 2022). Mitigating this problem is associated with conducting more recent and/or diverse data sets or the use of data augmentation procedures to artificially level the class distribution.

Limited Feature Set:

The data used to train the IDS's models might not include some significant elements that could get the models detecting smart cyber threats with higher success rate. Another potential direction for future research would be to embellish the detection and prediction methods with further network traffic statistics, including packet payload analysis, temporal correlations, and behavioural patterns. Improved feature representation allows inspecting the network at the deeper level and IDS models can identify atypical patterns that may have gone under of detection in the past.

Model Interpretability:

Models of machine learning, like logistic regression and random forest, may provide interpretability to some extent. Deep models of learning like convolutional neural networks (CNNs), however, often do not offer interpretability because of it complicated designing. Summarizing the amount of interpretability and, therefore, understanding of the rationale for IDS decisions and having a general view of the threat landscape are crucial. Next research study can be focused on the techniques which would boost the model interpretability, like various feature attribution methods, attention mechanisms, and model visualization ones.

Scalability and Resource Constraints:

Implementing of an IDS model in a resource restricted IoT/IIoT presents considerable issues like resource constrains and bandwidth limits. The future research should incorporate the development of lightweight and scalable IDS designs, which can operate very efficiently on edge devices, with less computational costs. The models can be compressed, quantized and hardware accelerated to speed up it deployment in constrained networks.

Adversarial Attacks:

With the growing trend of cyber threats, attackers might launch adversarial attacks aiming to spin detection by creating targeted IDS deception. Further research should examine the resilience of IDS models under adverse attacks and plot a way forward in countermeasure development. The following are examples of potential approaches that can be used to mitigator adversarial attacks on IDS performance: adversarial training, robust optimization techniques, and adversarial detection mechanisms.

Real-World Deployment and Validation:

To end, verifying the viability of IDS tools under real IoT and IIoT circumstances is crucial in order to demonstrate the utility and efficacy of these tools. The further research ought to include tests comprising of extensive field trials and deployment tests to ascertain the performance of IDS models under realistic operational conditions. Coordination with industry partners and involved actors of IoT system make it possible to open access to live data for conducting testing at a large scale.

Summing up, tackling down the weaknesses enumerated above, and following the recommended future research directions, one can considerably improve the level of intrusion detection in IoT and IIoT environments, which will eventually lead to increased resilience and protection of critical infrastructure from cyber threats.

4.7 Summary

As a result, the analysis of the data and the performance evaluation of all the ID models gave one an idea about various machine learning approaches' effectiveness in securing IoT and IIoT ecosystems against cyber threats. The exhibit of the descriptive analysis of the data set revealed the existence of different network traffic attributes and offered a generalized overview of attack types of distribution. What is more, the study of feature importance shows that some network features play the crucial role in distinguishing normal from abnormal traffic and, therefore, determines the design of effective intrusion detection systems.

The comparison of models showed the strong and weak sides of the unique ways it worked, therefore, demanding to achieve a compromise when considering the accuracy, interpretability, and computing efficiency. However, logic regression and random forest models, were transparent with the decision-making process and showed good performance across multiple metrics while CNNs were able to provide deeper insights into the complex network behaviour but lacked interpretability. It is also important that future work should be able to identify these trade-offs and try to find hybrid approaches that can compound the strengths of different models and yet also mitigate it weaknesses.

Although the obtained results are quite encouraging, there exist some limitations, such as data imbalance, insufficient feature set and large-scale issue (Alsaedi *et al.* 2020). These challenges require research activities that concentrate mainly on data augmentation, feature enrichment, and model optimization determining IoT and IIoT distinctive requirements. Additionally, to that, the improvement of intrusion detection model's interpretability and validating it effectiveness in real-world deployment environments need to be undertaken first as a way of ensuring it practical utility and effectiveness when fighting against emerging cyber-attacks.

To sum up, the link analysis provided in this study shows the need to develop a multi-pronged approach when detecting intrusions within the IoT and Industrial Internet of Things ecosystem. Through wards the right application of different types of machine learning methods, researchers and practitioners could benefit from the development of robust security solutions that are resistant to the ever-changing cyber threats while also reducing false positive and optimizing the utilization of the resources (Alsaedi *et al.* 2020). It is crucial to retain cooperation among academia, industry and governmental stakeholders, as well as ensure cybersecurity innovations in a world that is gradually connected (Abdel-Basset *et al.* 2020).

Chapter 5: Conclusion and Future Work

5.1 Conclusion

To sum up, this study demonstrates that intrusion detection system (IDS) plays a vital role in defeating emerging cyber threats for both IoT and IIoT environments. A thorough investigation of the dataset properties, model efficiency, the role of features, and comparisons has led to the attainment of meaningful results and conclusions concerning the handling of security breaches with machine learning algorithms. The descriptive approach to data analysis enabled the formulation of a basic understanding of the characteristics of network traffic and attack types in the IoT/IIoT networks. By this analysis the following investigation of the intrusion detection models effectiveness was initiated.

Amongst the variety of models one assessed, including logistic regression, random forest, and CNNs, these models showcased it different abilities in performing network traffic classification and detecting malicious activities. Performance of logistic regression and random forest was at a high level, with logistic regression providing great accuracy and random forest achieving high degree of feature importance. However, the CNNs participated, which imposed spatial connections over network data traffic for the purpose of identification of abnormal patterns. Although each model exhibited distinctive strengths and weaknesses, these models highlighted the fact that the primary factors that should be taken into consideration when developing an intrusion detection system are interpretability, accuracy, and computational efficiency.

Model comparison provided significant insights into the trade-offs related to different approaches and thereby gives directions for the future research that will seek the development of hybrid models that combine the best parts of several techniques. While logistic regression and random forest models are both easy to describe the decision-making processes and boast good performance in many indicators of performance, Convolutional neural networks provide deep analysis of complicated networks, but it is not so easy to interpret. Addressing data imbalance, refining feature sets, and improving model scalability through future research is the way to surmount the arising challenges that render impotent the intrusion detection systems in real-world as it is now.

Even though many advancements were achieved, some of the obstacles found are data augmentation, feature enrichment, along with the need for model optimization techniques that can be used on the specificity of IoT and IIoT landscapes. Also ensuring the implementation of the practical power of the intrusion detection systems show on real world deployment cases and the collaboration between academics, industry, and governmental stakeholders. Through this way, researchers and practitioners can beat the challenges and use the new digital tools and methodologies to develop credible security solutions that will outsmart the modern cyber threats and protect the critical infrastructure.

In summary, the study results stress the need of a comprehensive, multiple-method approach for intrusion detection of the IoT and IIoT environment. Through ongoing cooperation and innovation in cybersecurity, one can be up to date on the current nature of cyber threats and guarantee network or system integrity and security. Through the application of the findings from the research, the stakeholders can upgrade the resilience levels of the IoT and IIoT environments and lessen the chance of cyber-attacks affecting the society, economy, and national security.

5.2 Linking with objectives:

The outlined objectives serve as the guiding principles for the evaluation of machine learning (ML) and deep learning (DL) algorithms that are concerned with detection of intrusions within IoT and IIoT networks. Through periodic performance evaluations of these algorithms, researchers will seek to guarantee the security of interconnected systems and to maintain a level of cyber security. Scratching the surface, the first goal highlights the relevance of performance assessment of ML and DL techniques like logistic regression, random forest, and CNNs, in its ability to detect such illegal actions in IoT and IIoT scenarios. Scientists do this through multiple tests and evaluations that obtain knowledge about the strengths and weaknesses of various options. These gems will help to provide for choosing and using the intrusion detection systems.

Besides that, exploration of ensemble learning mechanisms like stacking integrated oversampling fits the second goal of enhancing ID accuracy. Via the combination of various models' strengths, and the data imbalance problems being handled, the researchers could bring the model accuracy to the higher level.

Researchers that investigate the behaviour of intrusion detection means in a variety resource condition can identify optimization decisions and design patterns for scaling up and more resource consumption.

The fourth goal deals with the performance of IDS mechanisms against developments in the cyber threats and assault methods targeted at smart environments. In the face of the continually developing cyber threat, which involves complex malware and zero-day exploits, it is imperative to consider the trace of intrusion detection systems in tracing and neutralizing the upcoming threats. Using the threat modelling and the simulation of the successful attack scenarios researchers can find flaws and shortcomings of the existing intrusion detection systems, which precedes the formation of the appropriate reactive pattern.

Lastly the fifth aim is also to provide useful information and recommendations to strengthen security designs and implementations in IoT and IIoT systems. Synthesizing the empirical findings and the real-world deployments will help the researcher to determine the best practices and guidelines for developing and deploying an intrusion detection system in a heterogeneous environment. These findings provide a basis for advocacy towards stakeholders like policy makers, system integrators, and end-users, with an eye on the prioritization of security concerns and the adoption of a comprehensive security outlook.

The overall objectives represent the crux of the study and guide moving forward the state of the art of intrusion detection for the IoT and IIoT networks the researchers will focus on the main research questions and will use an innovative methodology, to create strong security solutions that are intended to stop continuously growing cyber risks. Looking into the future, what should be done is to close the gap between theory research and practical implementation, leading to better scalability, efficiency, and effectiveness of intruder detection system in real world situations. Distributors of connectivity can promote continuous cooperation and exploitation to deal with the cybersecurity challenges associated with interconnectivity and the integrity and security of the IoT and IIoT ecosystems.

5.3 Recommendations:

Based on a detailed examination of intrusion detection on the Internet of Things (IoT) and Industrial Internet of Things (IIoT) networks, there are several suggestions, which should be the basis of future research and practice. The recommendations will attempt to solve central problems, optimise security capabilities, and ensure reliable operation of interrelated systems within the context of dynamic cyber threats landscape.

Firstly, the use of new machine learning (ML) and deep learning (DL) algorithms that are Geared precisely towards intrusion detection in IoT and IIoT structured environments should be continued. Scientists should be developing lightweight versions of the models, allowing it to run efficiently on devices with limited resources, while at the same time sustaining the accuracy of the detection of the abnormal activities. Besides, energy should be concentrated to investigate group methods for example stacking and integrated oversampling for improved insurances of the intrusion detection system.

In the same vein, one also has a role to play in doing empirical studies as well as deploying in the real world for the scalability and efficiency evaluation of intrusion detection techniques in varied IoT and industry 4.0 settings. Through comparison and analysis of different algorithms and assessing it performance under different resource constraints, researchers can make out effective strategies and assign some principles to the system to meet its scalability and resource utilization. On the other hand, the construction of standardized evaluation systems and datasets can advance the comparison and reproducibility between various studies.

Additionally, the attention should be also used for building in of proactive defences and threat intelligence feeds into the intrusion detection systems. With threat intelligence sources and the application of a dynamic rule-based solution, the industry can adapt to the new threats continuously and stop threats before it has it intended effect in real-time. Additionally, deployment of anomaly detection techniques as part of its repertoire of security measures will increase the range of intrusion detection systems that can be used in the detection of new and sophisticated kinds of attacks.

Besides, collaboration among the academic, industrial, and regulatory entities is indeed a priority to deal with the complex cybersecurity problems plaguing the cyber-physical and industrial internet ecosystems. Multidisciplinary initiatives and collaborative partnerships provide a platform to share knowledge, executing technology transfer, and the development of protocols and guidelines that are common to all. In addition to that, policymakers should give enough emphasis to cybersecurity issues and bring in the regulatory framework that will motivate the implementation of ‘secure-by-design’ principles and practices in the IoT and IIoT deployments.

Moreover, because of the necessity for the constant monitoring, auditing, and updating of intrusion detection systems to accustom with emergence of cyber threats and vectors of attack. Analyses that are carried out periodically detect the weaknesses potentially existing in runtime systems which can then be corrected by the application of remediation and security patches. Also, formulation of incident response protocols and establishing coordinated threat sharing platform would further boost the collective resilience of IOT and IIOT systems to cyber-attacks.

Finally, a regard should be given to user education and awareness efforts to enhance a security awareness culture among IoT and IIoT key holders. IoT product end-users, system integrators, and manufacturers must receive the proper security literacy about the risk factors of insecure IoT deployments and the imperative of putting in place impenetrable security measures. In addition, providing an incentive to implement system certifications and compliance standards can enable the participants to make security a priority issue in IoT and IIoT deployments.

Therefore, the complex cybersecurity issues exposed by IoT and IIoT systems necessitate a comprehensive solution from multiple angles including technology, regulation, collaboration, and education. Through the implementation of the recommendations, collectively all stakeholders can bring forth a higher level of security posture to intertwined systems and equally help in curtailing potential risks and sustain the resilience of IoT and IIoT ecosystems against growing cyber threats. Moving ahead, sustained work and collaboration between the organisations should be prioritized to guard the integrity, confidentiality, and availability of networks in the interconnected world.

5.4 Future Work

However, many paths for future research emerge moving on from intrusion detection of IoT and IIoT networks. Such future directions aim to fill current gaps, utilize emerging technologies, and elevate the robustness of these interconnected systems against fast-developing cyber threats.

One way ahead researching in this area could be the development of context-sensitive intrusion detection techniques which are customized for IoT and IIoT Environments. Intrusion detection can be improved by using contextual information like device type, location, and network behaviour. The system can understand policies that are based on normal and abnormal system activities enabling better and faster intrusion detection. Researchers can produce more accurate detection models that will reduce false alarms by including contextual knowledge. Thus, IIoT and IoT ecosystem's security level will be enhanced.

Moreover, the integration of anomaly detection methods and traditional signature-based detection techniques will lead to the formation of integrated and adaptive intrusion detection architecture. The methods of anomaly detection, e.g., machine learning-based anomaly detection and statistical analysis, can be utilized together with signature-based detection when detecting unusual and previously unseen attack patterns. By combining several detection methods, researchers can effectively build a multipurpose intrusion detection system that is capable of efficiently and accurately detecting against a broad arsenal of cyber threats.

In addition, further research resources should be allocated to building stronger Intrusion Detection Systems against techniques used in adversarial attacks and evasion tactics. Adversarial Machine learning for example has a huge impact on intrusion detection systems which can be overcome by manipulation or evasion by malware attacks. Investigating techniques of detection and mitigation of adversarial attacks, for instance, robust training methodologies and detection of adversarial examples may help to put the security of intrusion detection systems in good condition and make it more effective in real-world situations.

On the one hand, the evolution of edge computing and the blockchain ledger technology creates new perspectives for perceptual improvement in IoT and IIoT networks. Edge computing makes data processing and analysis possible near to where the data is generated, shortening the time required to get the data thus helping scalability in environments that are resource constrained. Researchers use edge computing platforms to develop ultralight intrusion detection solutions that are independent and operate effectively on edge devices to provide immediate threat detection and reaction.

One more, the inclusion of blockchain technology can strengthen the data integrity and immutability of intrusion detection logs and audit trails and keeps the authenticity of security events and consequent forensic analysis process in the situation of a security incident. Investigating new techniques in this field can be instrumental in designing protocols, which cover secure storing and managing intrusion detection data on blockchain networks, as well as mechanisms for protecting privacy and confidentiality in distributed networks.

Also, the goal for future works could be the development of intrusion detection system that is based on machine learning and is capable of self-adapting to changeable and evolving cyber threats. Through reinforcement learning methods and adaptive algorithms, researchers can build intrusion detection systems that are able to not only learn from new data but also adjust it detection mechanisms as its systems age. Such self-learning systems can overcome it shortcomings in the detection accuracy and adaptability against the newer attacks, thus obesifying the safety level of IoT and IIoT networks.

However, last but not least the empirical studies and implementations are necessary to ensure the efficiency and scalability of new IDI techniques in the actual IoT and IIoT environment. Large-scale study and on-site implementation can generate first-hand data on operational effectiveness, scalability, and practicality of IDS systems in the interest of resilient and high-performing protective means for interconnected systems and technologies.

In a nutshell, the future of intrusion detection in IoT and IIoT networks entails harnessing the power of the most recent technologies, building up detection capacities, and being ready for the changing digital landscape. The conduct of the above discussed research areas and overcoming of existing challenges will see scientists develop robust and adaptive IDS solutions that will in turn secure the integrity and safety of such systems in a world that is getting more and more interconnected.

References

Journals

Abdel-Basset, M., Chang, V., Hawash, H., Chakrabortty, R.K. and Ryan, M., 2020. Deep-IFS: Intrusion detection approach for industrial internet of things traffic in fog environment. IEEE Transactions on Industrial Informatics, 17(11), pp.7704-7715.

Albulayhi, K., Smadi, A.A., Sheldon, F.T. and Abercrombie, R.K., 2021. IoT intrusion detection taxonomy, reference architecture, and analyses. *Sensors*, *21*(19), p.6432.

Alsaedi, A., Moustafa, N., Tari, Z., Mahmood, A. and Anwar, A., 2020. TON_IoT telemetry dataset: A new generation dataset of IoT and IIoT for data-driven intrusion detection systems. Ieee Access, 8, pp.165130-165150.

Alsoufi, M.A., Razak, S., Siraj, M.M., Nafea, I., Ghaleb, F.A., Saeed, F. and Nasser, M., 2021. Anomaly-based intrusion detection systems in iot using deep learning: A systematic literature review. Applied sciences, 11(18), p.8383.
Altan, G., 2021. SecureDeepNet-IoT: A deep learning application for invasion detection in industrial Internet of things sensing systems. Transactions on Emerging Telecommunications Technologies, 32(4), p.e4228.

Attota, D.C., Mothukuri, V., Parizi, R.M. and Pouriyeh, S., 2021. An ensemble multi-view federated learning intrusion detection for IoT. *IEEE Access*, *9*, pp.117734-117745.
Awajan, A., 2023. A novel deep learning-based intrusion detection system for IOT networks. *Computers*, *12*(2), p.34.
Awotunde, J.B., Abiodun, K.M., Adeniyi, E.A., Folorunso, S.O. and Jimoh, R.G., 2021, November. A deep learning-based intrusion detection technique for a secured IoMT system. In International Conference on Informatics and Intelligent Applications (pp. 50-62). Cham: Springer International Publishing.

Baniasadi, S., Rostami, O., Martín, D. and Kaveh, M., 2022. A novel deep supervised learning-based approach for intrusion detection in IoT systems. *Sensors*, *22*(12), p.4459.
Bovenzi, G., Aceto, G., Ciuonzo, D., Persico, V. and Pescapé, A., 2020, December. A hierarchical hybrid intrusion detection approach in IoT scenarios. In *GLOBECOM 2020-2020 IEEE global communications conference* (pp. 1-7). IEEE.
Campos, E.M., Saura, P.F., González-Vidal, A., Hernández-Ramos, J.L., Bernabe, J.B., Baldini, G. and Skarmeta, A., 2022. Evaluating Federated Learning for intrusion detection in Internet of Things: Review and challenges. *Computer Networks*, *203*, p.108661.
De Souza, C.A., Westphall, C.B., Machado, R.B., Sobral, J.B.M. and dos Santos Vieira, G., 2020. Hybrid approach to intrusion detection in fog-based IoT environments. *Computer Networks*, *180*, p.107417.
Ellappan, V., Mahendran, A., Subramanian, M., Jotheeswaran, J., Khadidos, A.O., Khadidos, A.O. and Selvarajan, S., 2023. Sliding principal component and dynamic reward reinforcement learning based IIoT attack detection. Scientific Reports, 13(1), p.20843.

Friha, O., Ferrag, M.A., Benbouzid, M., Berghout, T., Kantarci, B. and Choo, K.K.R., 2023. 2DF-IDS: Decentralized and differentially private federated learning-based intrusion detection system for industrial IoT. Computers & Security, 127, p.103097.

Gaber, T., Awotunde, J.B., Folorunso, S.O., Ajagbe, S.A. and Eldesouky, E., 2023. Industrial internet of things intrusion detection method using machine learning and optimization techniques. Wireless Communications and Mobile Computing, 2023, pp.1-15.

Gad, A.R., Nashat, A.A. and Barkat, T.M., 2021. Intrusion detection system using machine learning for vehicular ad hoc networks based on ToN-IoT dataset. *IEEE Access*, *9*, pp.142206-142217.
Ge, M., Syed, N.F., Fu, X., Baig, Z. and Robles-Kelly, A., 2021. Towards a deep learning-driven intrusion detection approach for Internet of Things. *Computer Networks*, *186*, p.107784.
Hamouda, D., Ferrag, M.A., Benhamida, N. and Seridi, H., 2021, December. Intrusion detection systems for industrial internet of things: a survey. In 2021 International Conference on Theoretical and Applicative Aspects of Computer Science (ICTAACS) (pp. 1-8). IEEE.

Hindy, H., Bayne, E., Bures, M., Atkinson, R., Tachtatzis, C. and Bellekens, X., 2020, September. Machine learning based IoT intrusion detection system: An MQTT case study (MQTT-IoT-IDS2020 dataset). In *International Networking Conference* (pp. 73-84). Cham: Springer International Publishing.
Kasongo, S.M., 2021. An advanced intrusion detection system for IIoT based on GA and tree based algorithms. IEEE Access, 9, pp.113199-113212.

Khraisat, A. and Alazab, A., 2021. A critical review of intrusion detection systems in the internet of things: techniques, deployment strategy, validation strategy, attacks, public datasets and challenges. *Cybersecurity*, *4*, pp.1-27.
Lakshmanna, K., Kavitha, R., Geetha, B.T., Nanda, A.K., Radhakrishnan, A. and Kohar, R., 2022. Deep learning-based privacy-preserving data transmission scheme for clustered IIoT environment. Computational Intelligence and Neuroscience, 2022.

Le, T.T.H., Oktian, Y.E. and Kim, H., 2022. XGBoost for imbalanced multiclass classification-based industrial internet of things intrusion detection systems. Sustainability, 14(14), p.8707.

Liu, J., Yang, D., Lian, M. and Li, M., 2021. Research on intrusion detection based on particle swarm optimization in IoT. *IEEE Access*, *9*, pp.38254-38268. Liu, J., Yang, D., Lian, M. and Li, M., 2021. Research on intrusion detection based on particle swarm optimization in IoT. *IEEE Access*, *9*, pp.38254-38268.
Mohy-Eddine, M., Benkirane, S., Guezzaz, A. and Azrour, M., 2022. Random forest-based IDS for IIoT edge computing security using ensemble learning for dimensionality reduction. International Journal of Embedded Systems, 15(6), pp.467-474.

Mohy-eddine, M., Guezzaz, A., Benkirane, S. and Azrour, M., 2023. An effective intrusion detection approach based on ensemble learning for IIoT edge computing. Journal of Computer Virology and Hacking Techniques, 19(4), pp.469-481.

Nandanwar, H. and Katarya, R., 2024. Deep learning enabled intrusion detection system for Industrial IOT environment. Expert Systems with Applications, 249, p.123808.

Nimbalkar, P. and Kshirsagar, D., 2021. Feature selection for intrusion detection system in Internet-of-Things (IoT). *ICT Express*, *7*(2), pp.177-181.
Nuaimi, M., Fourati, L.C. and Hamed, B.B., 2023. Intelligent approaches toward intrusion detection systems for Industrial Internet of Things: A systematic comprehensive review. Journal of Network and Computer Applications, p.103637.

Qiu, H., Dong, T., Zhang, T., Lu, J., Memmi, G. and Qiu, M., 2020. Adversarial attacks against network intrusion detection in IoT systems. *IEEE Internet of Things Journal*, *8*(13), pp.10327-10335.
Sarhan, M., Layeghy, S., Moustafa, N., Gallagher, M. and Portmann, M., 2022. Feature extraction for machine learning-based intrusion detection in IoT networks. *Digital Communications and Networks*.
Sicato, J.C.S., Singh, S.K., Rathore, S. and Park, J.H., 2020. A comprehensive analyses of intrusion detection system for IoT environment. *Journal of Information Processing Systems*, *16*(4), pp.975-990.

Smys, S., Basar, A. and Wang, H., 2020. Hybrid intrusion detection system for internet of things (IoT). *Journal of ISMAC*, *2*(04), pp.190-199.
Soliman, S., Oudah, W. and Aljuhani, A., 2023. Deep learning-based intrusion detection approach for securing industrial Internet of Things. Alexandria Engineering Journal, 81, pp.371-383.

Tsimenidis, S., Lagkas, T. and Rantos, K., 2022. Deep learning in IoT intrusion detection. *Journal of network and systems management*, *30*(1), p.8.
Wahab, O.A., 2022. Intrusion detection in the iot under data and concept drifts: Online deep learning approach. *IEEE Internet of Things Journal*, *9*(20), pp.19706-19716.
Websites

mdpi.com, 2024. Available at: https://www.mdpi.com/2073-431X/12/2/34 [Accessed on: 02.05.24]

researchgate.net, 2024. Methodology Flowchart. Accessed from: <https://www.researchgate.net/figure/Methodology-flowchart_fig1_341072739> Accessed on: 07.05.2024

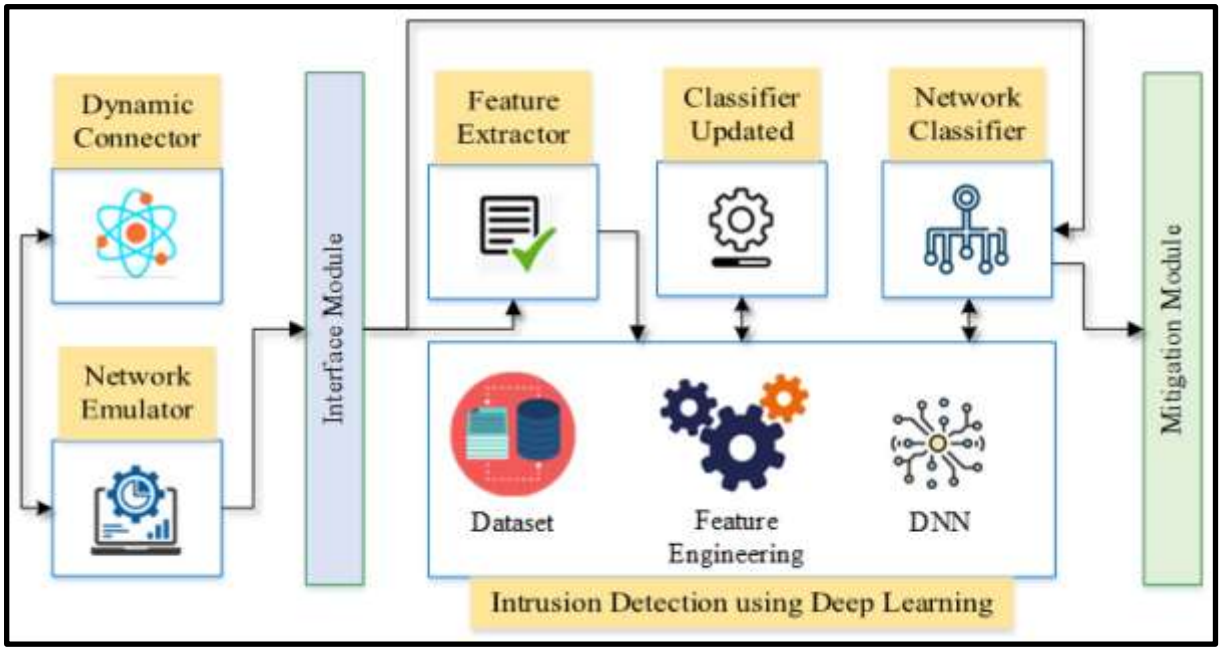
springer.com, 2024. Available at: https://link.springer.com/article/10.1007/s10586-022-03810-0 [Accessed on: 11.05.24].

Appendix:

Dataset link: Edge-IIoTset Cyber Security Dataset of IoT & IIoT (kaggle.com)

Abbreviations:

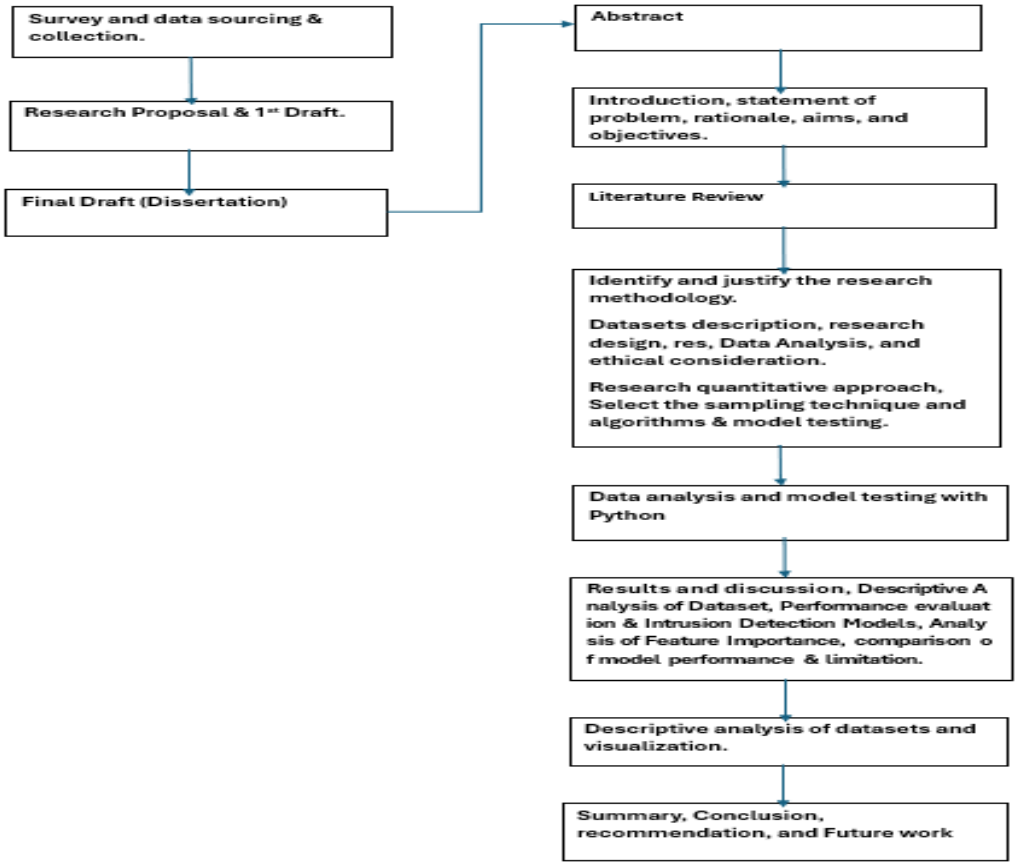
1. IIoT: Industrial Internet of Things
2. Edge: Edge Computing
3. Python: Programming language commonly used for data analysis and software development.
4. RapidMiner: Data science platform for analytics, machine learning, and artificial intelligence.
5. GDPR: General Data Protection Regulation, European Union regulation for data protection and privacy.
6. AI: Artificial Intelligence
7. ML: Machine Learning
8. CSV: Comma-Separated Values, a file format used to store tabular data.
9. IoT: Internet of Things
10. SSL: Secure Sockets Layer, a cryptographic protocol for securing communications over a computer network.

11. 

12. Figure 1: Efficacy of Machine Learning and Deep Learning Algorithms in Intrusion Detection

13. (Source: mdpi.com, 2024)

Flow Chart:



Figures 3.1: Methodology Flowchart
(Source: self-created)

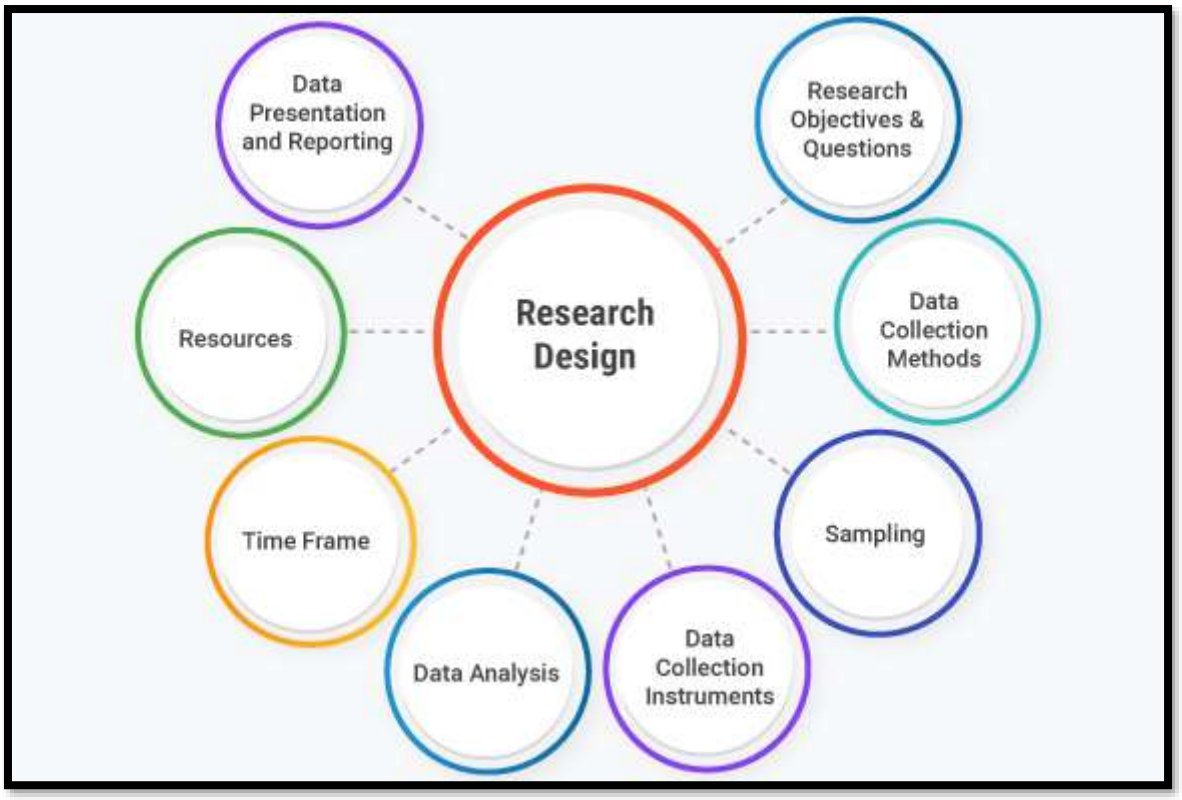
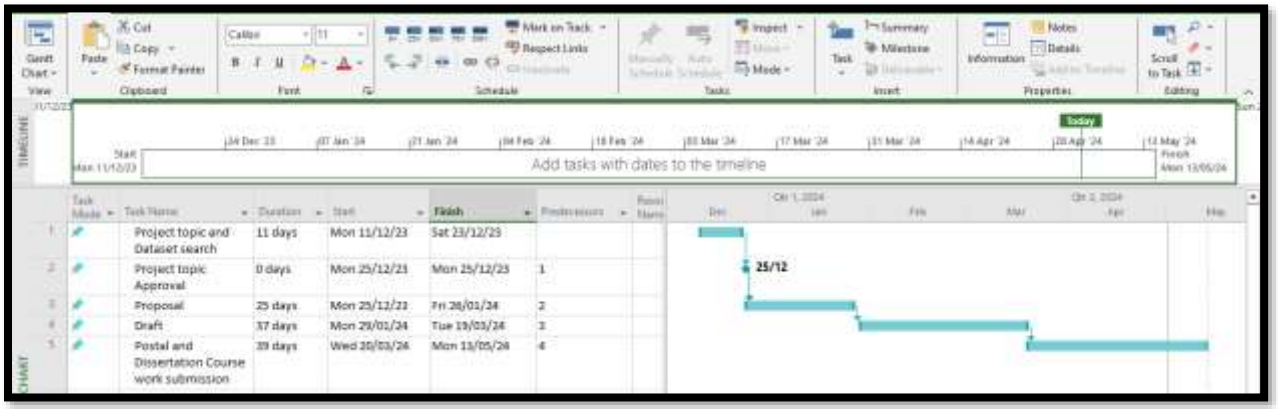


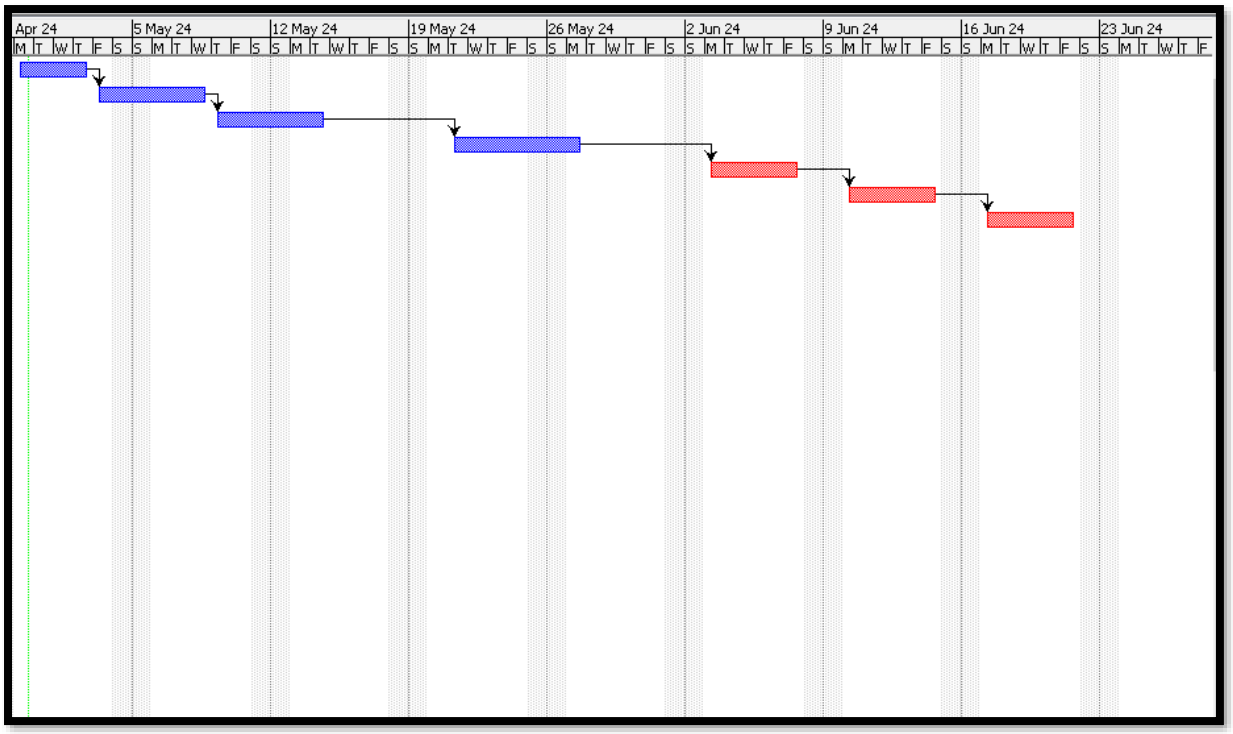
Figure 3.4.1: Research Design
(Source: researchgate.net, 2024)

Project Plan



Initial Gantt-Chart

(Source: self-created)



Final Gantt-Chart

(Source: self-created)

Python Codes:

Logistic regression accuracy:

```
In [26]: # Calculating accuracy
         accuracy_lr = accuracy_score(y_test, y_pred_lr)
         print("Accuracy:", accuracy_lr)

         Accuracy: 0.8592839036755386
```

CNN Model Implementation

In [40]:
```
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Flatten, Conv1D, MaxPooling1D, Dropout
```

```
WARNING:tensorflow:From C:\Users\User\AppData\Roaming\Python\Python311\site-packages\keras\src\losses.py:2976: The name tf.loss
es.sparse_softmax_cross_entropy is deprecated. Please use tf.compat.v1.losses.sparse_softmax_cross_entropy instead.
```

In [41]:
```
# Reshape X to a 3D array suitable for CNN
X = np.array(X).reshape(X.shape[0], X.shape[1], 1)
```

In [42]:
```
# Splitting the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

In [44]:
```
# Reshape X_train and X_test to 2D arrays for scaling
X_train_2d = X_train.reshape(X_train.shape[0], -1)
X_test_2d = X_test.reshape(X_test.shape[0], -1)
```

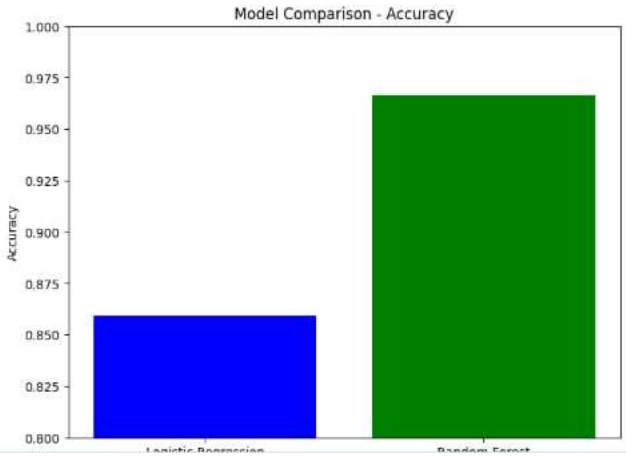
In [45]:
```
# Standardize features
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train_2d)
X_test_scaled = scaler.transform(X_test_2d)
```

In [46]:
```
# Building the CNN model
model = Sequential([
    Conv1D(filters=32, kernel_size=3, activation='relu', input_shape=(X_train.shape[1], 1)),
    MaxPooling1D(pool_size=2),
    Conv1D(filters=64, kernel_size=3, activation='relu'),
    MaxPooling1D(pool_size=2),
    Flatten(),
    Dense(128, activation='relu'),
    Dropout(0.5),
    Dense(1, activation='sigmoid')
])
```

```
WARNING:tensorflow:From C:\Users\User\AppData\Roaming\Python\Python311\site-packages\keras\src\backend.py:873: The name tf.get_
default_graph is deprecated. Please use tf.compat.v1.get_default_graph instead.

WARNING:tensorflow:From C:\Users\User\AppData\Roaming\Python\Python311\site-packages\keras\src\backend.py:6642: The name tf.nn.
max_pool is deprecated. Please use tf.nn.max_pool2d instead.
```

Model Comparison Plot

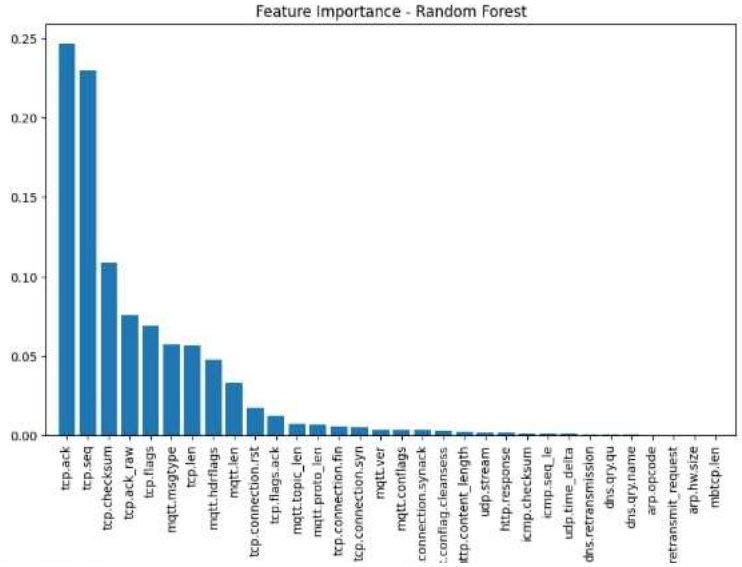
In [39]:
```
import matplotlib.pyplot as plt

# List of model names
models = ['Logistic Regression', 'Random Forest']

# List of accuracies
accuracies = [accuracy_lr, accuracy_rf]

# Plotting the comparison
plt.figure(figsize=(8, 6))
plt.bar(models, accuracies, color=['blue', 'green'])
plt.title('Model Comparison - Accuracy')
plt.xlabel('Models')
plt.ylabel('Accuracy')
plt.ylim(0.8, 1.0)  # Set y-axis limit for better visualization
plt.show()
```



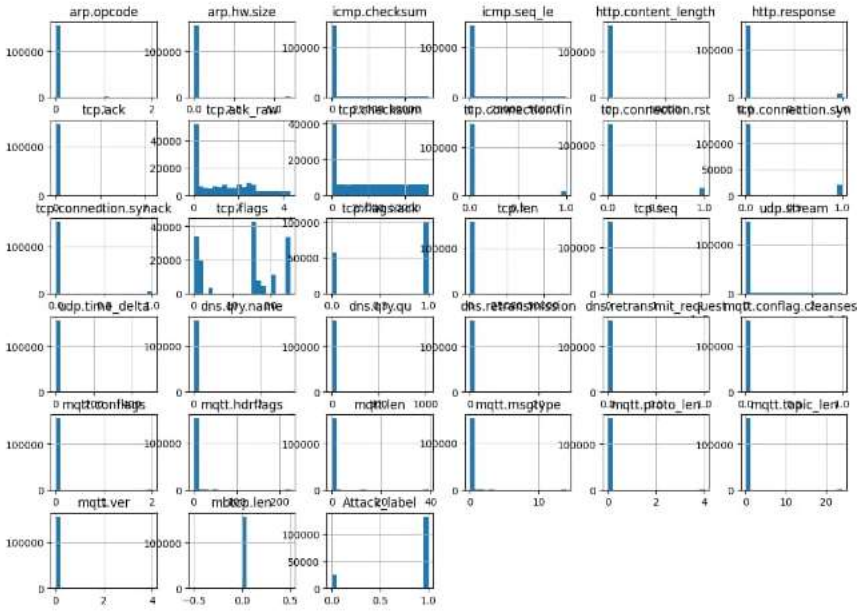
```
In [27]: # Generating confusion matrix
         conf_matrix_lr = confusion_matrix(y_test, y_pred_lr)
         print("Confusion Matrix:")
         print(conf_matrix_lr)

         Confusion Matrix:
         [[  544  4441]
          [    0 26575]]
```

Feature Importance Plot

```
In [37]: # Plotting feature importance
         plt.figure(figsize=(10, 6))
         importances = rf_model.feature_importances_
         indices = np.argsort(importances)[::-1]
         plt.bar(range(X.shape[1]), importances[indices], align="center")
         plt.title("Feature Importance - Random Forest")
         plt.xticks(range(X.shape[1]), X.columns[indices], rotation=90)
         plt.xlim([-1, X.shape[1]])
         plt.show()
```



Histogram



Random Forest Classifier Implementation

In [31]:
```
from sklearn.ensemble import RandomForestClassifier
```

In [32]:
```
# Creating Random Forest model
rf_model = RandomForestClassifier()

# Training the Random Forest model
rf_model.fit(X_train, y_train)
```

Out[32]:
```
RandomForestClassifier
RandomForestClassifier()
```

In [33]:
```
# Making predictions
y_pred_rf = rf_model.predict(X_test)
```

In [34]:
```
# Calculating accuracy
accuracy_rf = accuracy_score(y_test, y_pred_rf)
print("Accuracy:", accuracy_rf)
```

```
Accuracy: 0.9666032953105197
```

In [35]:
```
# Generating confusion matrix
conf_matrix_rf = confusion_matrix(y_test, y_pred_rf)
print("Confusion Matrix:")
print(conf_matrix_rf)
```

```
Confusion Matrix:
[[ 4312   673]
 [  381 26194]]
```

In [49]:
```
# Evaluating the model
test_loss, test_accuracy = model.evaluate(X_test, y_test)
print("Test Loss:", test_loss)
print("Test Accuracy:", test_accuracy)
```

```
987/987 [==============================] - 2s 2ms/step - loss: 1.0102 - accuracy: 0.8420
Test Loss: 1.0101573467254639
Test Accuracy: 0.8420469164848328
```

Visualization - 1

In [16]:
```
# Visualization 4: Bar Chart for Attack Types
plt.figure(figsize=(12, 8))
sns.countplot(y='Attack_type', data=df, order = df['Attack_type'].value_counts().index)
plt.title('Count of Attack Types')
plt.xlabel('Count')
plt.ylabel('Attack Type')
plt.show()
```