A dissertation submitted to the **University of Greenwich** in partial fulfilment of the requirements for the Degree of Master of Scienc *in*

**Big Data & Business Intelligence**

**Adaptive encryption scheme for IoT sensors network**

**Name:** XYZ

**Student ID:** 000

**Supervisor:** XYZ

**Submission Date:** August, 2024

**Word Count:** 17845

Adaptive encryption scheme for IoT sensors network

XXXXX

Computing & Mathematical Sciences, University of Greenwich, 30 Park Row, Greenwich, UK

*(Submitted 10 August 2024)*

# Abstract

IoT is a new complex paradigm of different interconnected machines that share tiny amount of resources and have limited computational power therefore the encryption method offered in this thesis is an adaptive IoT sensor protection approach for the respective IoT conditions and constraints. The proposed Scheme has incorporated the concept of machine learning for the classification of the data based on criticality, XOR encryption for low critical data to save transaction overheads while HE for highly critical ones for stringent secure transactions. To evaluate the scheme, a large number of experiments were performed with the help of the provided N-BaIoT dataset to show that the scheme successfully regulates the trade-off between the security level and computation complexity. In its result, it is evidenced that the adaptive encryption approach not only makes the data security reliable But also the network functional availability, which is vital for IoT. This research provides a new, adaptive approach to solving IoT security, that may have usage in a plethora of industries as the sensitivity of data and the constraint of resources increases. Further work is recommended for betterment of the classification models, generalization of the scheme, and for studying the interaction of the proposed scheme with other security frameworks like block chain to make Iot networks more secure.

**Keywords:** Adaptive Encryption; IoT Security; Machine Learning; XOR Encryption; Cybersecurity.

# Acknowledgements

I would especially like to thank XXXXXXXXXX for agreeing to be my supervisor and for his consistent advice, feedback, guidance and support throughout the lifecycle of this Msc data warehouse project.

I want to thank both XXXXXXXXXXXXX and XXXXXXXXX for agreeing to have the project demonstration on the schedule day.

# Table of Contents

[Abstract i](#_Toc175229499)

[Acknowledgements ii](#_Toc175229500)

[Table of Contents iii](#_Toc175229501)

[Chapter 1 Introduction 7](#_Toc175229502)

[1.1 Background 7](#_Toc175229503)

[1.2 Motivation 11](#_Toc175229504)

[1.3 Aims and Objectives 11](#_Toc175229505)

[1. 4 Limitation of the Study 11](#_Toc175229506)

[1. 5 Implication of the Study 12](#_Toc175229507)

[1. 6 Thesis Structure 12](#_Toc175229508)

[1. 7 Conclusion 13](#_Toc175229509)

[Chapter 2 Literature Review 14](#_Toc175229510)

[2.1 Privacy and User Behaviour 14](#_Toc175229511)

[2.1.1 Privacy 14](#_Toc175229512)

[2.1.2 User Behaviour 15](#_Toc175229513)

[2.2 Privacy in IoT and AI Environment 17](#_Toc175229514)

[2.3 Key Enabling PETs Technologies 21](#_Toc175229515)

[2.3.1 Anonymisation 22](#_Toc175229516)

[2.3.2 Differential privacy 23](#_Toc175229517)

[2.3.3 Secure Computing 24](#_Toc175229518)

[2.3.4 Federated Learning 28](#_Toc175229519)

[2.4 PETs Challenges 30](#_Toc175229520)

[2.4.1 Challenges in DP 30](#_Toc175229521)

[2.4.3 Challenges in Secure Computation 30](#_Toc175229522)

[2.4.4 Challenges in FL 32](#_Toc175229523)

[2.5 Privacy Challenges in IoT and AI Environment 32](#_Toc175229524)

[2.5.1 IoT Challenges 32](#_Toc175229525)

[2.5.2 AI Challenges 34](#_Toc175229526)

[2.6 Privacy and GDPR 35](#_Toc175229527)

[Chapter 3: Research Methodology and Implementation 38](#_Toc175229528)

[3.1 Research Design 38](#_Toc175229529)

[3.2 Data Sources and Collection 39](#_Toc175229530)

[3.3 Quantitative Methods 39](#_Toc175229531)

[3. 4 Qualitative Methods 40](#_Toc175229532)

[3.5 Research Implementation 41](#_Toc175229533)

[3.5.1 Dataset Preparation 41](#_Toc175229534)

[3.5.2 Model Training 43](#_Toc175229535)

[3.5.3 Adaptive Encryption 43](#_Toc175229536)

[3.5.4 Validation Process 45](#_Toc175229537)

[3.6 Conclusion 46](#_Toc175229538)

[Chapter 4 Results and Discussion 46](#_Toc175229539)

[4.1 Model Performance Evaluation 46](#_Toc175229540)

[4.1.1 Confusion Matrix for results analysis 47](#_Toc175229541)

[4.1.2 ROC Curve and AUC Analysis 48](#_Toc175229542)

[4.2 Discussion of Feature Importance 49](#_Toc175229543)

[4.3 Generalization Of Their Model And The Likelihood Of Over-Fitting 50](#_Toc175229544)

[4.4 Comparison of the Model with Other Models 50](#_Toc175229545)

[4.5 Recommendations and Implementation 50](#_Toc175229546)

[4.6 Limitation of the Present Study and Possible Directions for Further Research 51](#_Toc175229547)

[4.7 Conclusions 51](#_Toc175229548)

[Chapter 5 Conclusion and Future Works 52](#_Toc175229549)

[5. 1 Introduction 52](#_Toc175229550)

[5. 2 Conclusion 52](#_Toc175229551)

[5. 3 Future Work 52](#_Toc175229552)

[References 53](#_Toc175229553)

List of Figures

[Figure 1 Research Architecture for Adaptive Encryption Scheme 36](#_Toc175229083)

[Figure 2 Features importance 40](#_Toc175229084)

[Figure 3 Data visualization before training and testing 41](#_Toc175229085)

[Figure 4 ML model output 41](#_Toc175229086)

[Figure 5 Adaptive Encryption output 42](#_Toc175229087)

[Figure 6 validation output 43](#_Toc175229088)

[Figure 7 Confusion matrix 46](#_Toc175229089)

[Figure 8 ROC curve analysis 47](#_Toc175229090)

List of Tables

# Chapter 1 Introduction

### 1.1 Background

Over the last decade, Artificial Intelligence (AI) and the Internet of Things (IoT) have advanced significantly, changing various industries such as healthcare, finance, and smart cities. The vast and complex network of connected smart devices forms the IoT ecosystem, through which large volumes of data are generated, processed, transmitted, and stored, thus enabling extremely high levels of automation and intelligence. Several sectors are transformed due to IoT applications, improving efficiency, convenience, and connectivity. For instance, smart homes and wearable devices provide personalised services in daily life and function as digital housekeepers, healthcare providers, and fitness coaches (Li and Palanisamy 2019). In addition, other applications of the IoT in smart buildings and smart cities are increasingly focused on environmental awareness and user convenience. The affordable availability of smart things, together with high-speed wireless Internet, drives the growth of the IoT and allows physical things to communicate and cooperate through different communication protocols (ibid). IoT systems are converting our lives to be more automatic, intelligent, and interconnected, improving the levels of convenience and efficiency, thus moving the quality of living forward.

AI is a branch of computer science defined as the machines that perform activities and replicate human intelligence. They learn from a set of data and make decisions or predictions. Such activities include learning, reasoning, problem-solving and understanding a natural language. This includes general AI, which replicates the human mind to perform various tasks in different realms and narrow AI, which is mainly oriented to one specific task with high precision but limited scope (Al-Rubaie and J. Morris 2019).

The system’s ability to learn from data automatically, without being programmed through explicit instructions, is known as Machine Learning (ML), an AI-driven science. There are two main categories of ML algorithms, supervised learning and unsupervised learning. In supervised learning, through trained labelled data, an accurate model should be built toward the ability to predict outcomes with new data. Unsupervised learning is the process of trying to learn such patterns in a dataset without providing output labels (Meurisch and Mühlhäuser 2021). For example, a number of technologies support the personalisation of end users' experience with AI services, such as IBM Watson Health, which can offer treatment recommendations based on patient data. Another example would be Bank of America's Erica, which offers personalised advice on banking, with insights into spending based on real-time data processing and low-latency interactions with nearby IoT devices (ibid).

Among the wide variety of techniques that are used within AI systems, Logistic Regression is a linear model widely applied to binary classification tasks. It estimates the probability that an input belongs to a class. These models are based on the logistic function, which can take any real-valued number and map it to a value between 0 and 1 (Menard 2002). It's used to work on tasks whose intent is to predict probabilities. The most basic idea in this is that it models the relationship between features (input variables) and a target variable, as the linear combination of the input features, weighted by coefficients.

This revolution in various industries has been driven by the ability to collect, process, and analyse large volumes of data in real-time; increasing efficiency and automation, however, may lead to privacy challenges. The general availability of low-cost sensing and communication technologies development within the IoT system helps create smart services that rely on the generated data by embedded devices in everyday life (Li and Palanisamy 2019). These developments significantly raise the value of raw data, make the extraction of patterns feasible, and allow for greater personalisation and support for end users via AI applications (Ghayyur et al. 2020). Sharing data with third parties is another critical challenge in data privacy. Even with the massive deployment of IoT, many organisations do not actively monitor the related third-party access privacy risks. Effective management of personal data is vital in building trust and compliance with various regulations.

Another security issue is that attacks on IoT devices can compromise unauthorised access to personal and confidential data, thus risking data privacy. In other words, interception of data transmissions and package tampering of the IoT device led to sensitive information leakage and a data privacy compromise (Meneghello et al. 2019). Moreover, substantial privacy concerns arise with AI services that are fundamentally dependent on users’ data. Behavioural pattern collection from users and devices forms a foundation base in AI and big data analytics applications (Ghayyur et al. 2020). Most AI applications process local computation, which introduces significant problems with data security. In ML, an adversary can learn some private knowledge regarding the training data of a model. For example, membership inference attacks are privacy attacks when an adversary can exploit whether some of the specific data samples had been used in the model's training process (Farokhi and Kâafar 2020).

Additionally, user behaviour is important in protecting personal data and ensuring privacy in the digital era. IoT and AI are used for increased convenience and efficiency, although they raise severe privacy and security issues. Users are heavily dependent on these AI systems, and that results in a lack of attention to data collection practices and the unconscious sharing of sensitive information (Zheng et al. 2018). However, many users are ignorant that AI may analyse and infer sensitive information from their IoT data, causing privacy infringement. Furthermore, it is common for users to forget to set the right privacy level by using default settings or common weak passwords (Tawalbeh et al. 2020). There is also the issue of inherent trust in the device just because it belongs to a well-known brand (Zheng et al. 2018). Such behaviours highlight the need to educate users about the threatening behaviour that leads to privacy (Nguyen and Vu 2023).

Regulations such as the General Data Protection Regulation (GDPR) of Europe and the California Consumer Privacy Act (CCPA)—insist that the data controller must have strong privacy-preserving mechanisms that protect and secure users’ data. To comply with privacy laws such as CCPA and GDPR and ensure business continuity, there is a need to reduce the collection of sensitive personal data as a measure to prevent identity theft and moderate the effect of data breaches. Moreover, allowing users to be in control of their data and giving them the right to reject sharing their data. More importantly, privacy ensures that human rights regarding freedom of expression and will are observed. These practices are vital in the organizations and companies that handle users' data to earn user trust (Ghayyur et al. 2020).

Further, advanced privacy-preserving techniques (PETs) have been developed in the past decade to ensure the protection of sensitive data and maintain user trust (Abadie 2021). PETs are technical measures that protect the user identity through the possibilities of anonymity, pseudonymity, unlikability, and unobservability of the user and the data subject (Heurix et al. 2015; Pfitzmann and Hansen 2006). They enable entities to cooperate in extracting insights from confidential information without revealing users’ sensitive information, allowing computations in an untrusted environment based on trustworthy information. This includes various techniques, such as anonymity, in which the elements that make a person identifiable, such as names and addresses, are carefully removed. Differential Privacy (DP) depends on adding noise randomly to the results of a query and using Secure Multiparty Computation (SMPC) allows multiple parties to process data without revealing individual inputs. Other privacy preservation methods include homomorphic encryption (HE), which operates mathematical calculations on encrypted data, and federated learning, which involves training models in a decentralised method that does not need to publicise the raw data. According to Al-Rubaie and J. Morris (2019), these strategies are critical for protecting users’ privacy and maximising the benefits of IoT, cloud computing, and AI while ensuring users’ privacy.

This report will describe the importance of privacy for IoT and AI from a technical and personal point of view, explain existing concerns in data privacy regarding the AI and IoT landscape; and then introduce current advancements in PETs and their use within AI and the IoT. Moreover, investigates how sensitive information can be effectively protected by training an ML model to classify data sensitivity into low and high criticality classes. This research will use the N-BaIoT dataset (Meidan et al. 2018), based on the features extracted from the packets, such as size, communication protocol, transmission frequency, time transmission, and port numbers. This classification notably minimises the overhead computation, thus ensuring that only the necessary level of security must be implemented based on sensitivity. For very low critical data, light XOR encryption ensures basic security without compromising system performance. On the other hand, a homomorphic encryption-like method is implemented to the highest degree of criticality data, as it provides a strong computational method that securely processes and analyses data without exposing the information it carries. This two-layered approach toward encryption ensures that a balance is created between the security and efficiency of the data in AI and IoT applications by handling the specific needs of the sensitivity level of the data. The evaluation of a Logistic Regression model's ability to classify network traffic data as Low or High criticality, achieving a high accuracy of 99.97% during the testing phase. The dataset was split into 80% for training and 20% for testing, with features normalized to ensure that no single feature disproportionately influenced the model. The confusion matrix revealed that the model correctly identified 24,563 instances of high-critical data (True Positives) and 9,852 instances of low-critical data (True Negatives), with a False Positive rate of zero, indicating that no low-critical data was misclassified as high-critical. Additionally, the False Negative count was minimal at 10, meaning almost all high-critical data was accurately classified. These results highlight the model's effectiveness and reliability in distinguishing between different types of network traffic, making it highly suitable for practical applications where precise classification is essential.

## 1.2 Motivation

As AI and IoT technology develop, protecting user privacy is critical. Integrating AI and IoT brings several benefits but involves several data privacy and security risks. The enormous volume of data usually flows through these technologies, most of which carry sensitive personal information, making them a potential target for malicious attacks and privacy breaches. Several tools and techniques are provided by PETs, guaranteeing the confidentiality and security of data with the benefits derived from AI and IoT applications. Looking at the vast volume of data generated by AI and IoT devices, it is crucial to determine the criticality of this data to apply appropriate security measures. Implementing security based on data sensitivity can optimise system performance and resource utilisation. Most of the PET techniques face computational overhead as a cost for the higher level of security that they provide. Light XOR encryption is suitable for low-critical data, offering basic security without compromising efficiency. The Homomorphic encryption-like method ensures data security for high-critical data by allowing mathematical computations on encrypted data, therefore maintaining privacy.

## 1.3 Aims and Objectives

The main objective of this dissertation is to explore and design an ML algorithm to determine the criticality of the data that can be well-integrated into AI and IoT systems.

1. To research the critical role of privacy in AI and IoT from technical and personal perspectives and to address the increasing concern toward data privacy in these fields.
2. Explore and analyse the current solutions of PETs and their applications in AI and IoT.
3. Evaluate the challenges in implementing PETs in AI and IoT.
4. Design a model that classifies the data into low or high criticality, using PEPs to reduce computational overhead while maintaining security.
5. Analyse how PETs can assist in aligning with privacy regulations such as GDPR.

## 1. 4 Limitation of the Study

Consequently, the following section will outline the scope of the study with particular emphasis on the particular areas of artificial intelligence and IoT security that the dissertation will cover. It will also highlight the limitations of the study, it is with regards to the technologies, the datasets adopted, and encryption methods used.

## 1. 5 Implication of the Study

In this section, the focus will lie on the significance of its findings for the general field of cybersecurity and data privacy. Finally, it will describe in detail how the discovering of this dissertation could be useful in improving the security of sensors in IoT sensor networks and how the outcomes will shape the direction for the design of better PETs in Artificial Intelligence Systems.

## 1. 6 Thesis Structure

As postulated in the earlier chapters of this research thesis proposal, this format of developing the research questions and objectives has been made in a manner that there will be maximum organization of the various sections of the research study.. The thesis is divided into the following chapters:The thesis is continued in the following chapters:

Chapter One: Background to the Study – The first chapter of the study concerns the background of the research work to be conducted and the objectives of the study apart from other essential aspects of the study’s feasibility and importance. This gives an indication of the coverage of the study, and the objectives of the dissertation to be undertaken.

Chapter 2: Literature Review: In this particular chapter, papers that relate to; AI, IoT and the potentially enhancing technologies addressing the privacy issues are presented and discussed. This chapter has looked at the extent and the ways and methods that have been offered and presented with the view of protecting IoT Sensor networks and also the Position of PETs in AI and IoT systems. The role of the review chapter of this study is therefore to prepare the stage of the study in terms of establishing areas of research that require further work.

Chapter Three: Methodology – In this chapter, the author positions the method applied for selecting datasets, appropriate technologies, as well as encryption techniques in the course of the study. Regarding the rules of the ML algorithm, the techniques from which they originated and the mutual application of the ML algorithm and cryptography and experiments and grading.

Chapter 5: Conclusion – Total conclusion of the carried out experiments and their results are presented in this chapter. In addition, the live demonstration of the data classification employing the proposed ML model, along with the performance analysis of the proposed adaptive encryption scheme is also presented. As she in previous chapters, this chapter also worries general conclusions of the book for safety of AI and IoT.

Chapter 5: Conclusion and Future Works – This uses and brings conclusion to the work stated in. future works is a section of the study that outlines the research propositions that is the focal assertion of the research and its implications. With it, IE provides the limitations of novel studies and directions of subsequent work for improving the security and dependability of the IoT sensor network.

## 1. 7 Conclusion

The conclusion section of this research work offers a summary of the above study in line with the implementation of an adaptive encryption technique that offers security in IoT sensor networks without placing much pressure on the computation resources. The study establishes privacy enhancing technologies (PETs) as essential mechanisms of protecting data in AI and IoT systems: it also shows the potential of advanced machine learning algorithms for classifying data according to its sensitivity. The strength of the study is also the subject of the conclusion, and the authors announce their achievements and the difficulties they faced during the work, namely the high integration complexity of PETs and the need for further research. The ideas presented in this thesis are relevant to the development of cybersecurity as a whole, along with the particular developments of IoT protection.

# Chapter 2 Literature Review

## 2.1 Privacy and User Behaviour

### 2.1.1 Privacy

The unprecedented data generation, collection, and processing scale has led to significant improvements in AI and IoT. Concerns about data privacy and security have been rapidly increasing. Privacy is not just a personal concern; it's an issue related to individual freedom, dignity, and security (Abdul and Seong Oun 2023). Support individuals in managing how their personal data is handled and making informed decisions about how it is used; this will help to protect themselves from identity theft, financial fraud, discrimination, and invasive surveillance. The importance of data privacy is magnified in the context of the IoT and AI. Users' habits and routines are collected through IoT devices embedded in our daily lives; this highlights the need for strong privacy protection to ensure data misuse or unauthorised disclosure. Additionally, AI systems used for decision-making and predictions depend heavily on personal data, which calls for strict privacy safeguards to avoid bias and support fairness.

Data protection is the foundation of trust in digital technologies, which drive innovations and economic growth. Strong privacy protection is required to ensure the security of systems and resilience of our interconnected systems. A data breach can trigger far-reaching effects, including cyber-attacks, financial loss, or consumer confidence loss. These attacks on privacy have, beyond personal concerns, huge security implications. They highlight the basic need for trust in entities responsible for safeguarding personal data and ensuring its security and protection (Heurix et al. 2015).

Even with the widely familiar concept, privacy remains hard to define. It is a multifaceted, complex idea that may be viewed from legal, philosophical, or technical perspectives. In 1890, Warren and Brandeis were among the first to recognise this emergent privacy challenge posed by upcoming technologies associated with societal transformation. They created the concept of “the right to privacy” as a general perception to shield individuals from the unwarranted revelation of personal information, including individuals’ private thoughts, feelings, and other intimate matters (Warren and Brandeis 1890) cited in (Meurisch and Mühlhäuser 2021). A more focused term is information privacy, which relates to the processing and disclosing of personal information. Westin (1968). regards privacy as a right "to control, edit, manage, and delete information about oneself and to determine when, how, and to what extent that information is communicated to others" (ibid). This includes the right of an individual to control how their personal information is handled and shared with others. Privacy is one of the fundamental human rights guaranteed under Article 8 of the European Convention on Human Rights and Fundamental Freedoms (Council of Europe 2010).

Abdul and Seong Oun (2023) stated that privacy could be divided into four significant dimensions: information, communication, territory, and the body. According to them, information privacy is related to personal data–how it is collected, stored, shared, and used. It can be considered part of information security, which is the combination of the three fundamental attributes called CIA: confidentiality, integrity, and availability (Avizienis et al. 2004). Communication privacy maintains the confidentiality of messages, preventing unauthorised access to the data. Territorial privacy protects physical location and space. Bodily privacy conceals the body and its information from others and intrusive procedures.

Personal data can be stored in different forms, each suitable for different uses. For example, time series record values of a variable at successive points in time, such as daily stock prices. Traces record sequences of activities, as with the history of web browsing. Matrices represent data using grids, useful in computations or images such as pixel intensity values in a grayscale image. These formats improve the effective analysis and utilisation of personal data (Abdul and Seong Oun 2023).

### 2.1.2 User Behaviour

The incorporation of IoT and AI in our daily lives and real-time applications is revolutionising the processing of data handling. These developments increased the convenience and efficacy of everyday activities; however, they raise concerns over privacy and security issues. The first defence line for protecting personal data and maintaining privacy is user behaviours with these technologies. Users often rely heavily on AI-based automated systems unknown the data processing and collection practices. This results in overreliance on these systems and probably the extensive sharing of information, unaware of the risk associated with it (Zheng et al. 2018). Most users do not know about the intensive data collection that IoT devices are capable of. For example, Google Smart Assistant (GSA) has placed significant privacy risks due to its unusual requests for permissions and utilisation of sensitive APIs that might compromise user data security (Elahi et al. 2019). Users might have no idea what these permissions are for, and allowing such permissions can lead to misuse or overreach in the data collection. For instance, the application asks for permission to access the microphone and location; this indicates that the application could listen at any time to what is happening without the explicit consent or knowledge of the user.

Additionally, in order to access a service, users consent to the privacy statements without reading what they agree to due to the complexity and length of these statements. This action threatens their privacy as they may have consented to use their personal data under terms and conditions. Repurposing of the data in AI analyses may also need to be re-secured with informed consent, posing further privacy threats (Zheng et al. 2018). Nguyen and Vu (2023) state that the need for more transparency in communication and implementation of these practices is challenging. Since users cannot evaluate associated risks, leaving their privacy and safety exposed.

Users may need to realise that data, non-audio/visuals, could be used by ML algorithms to infer sensitive information such as sleep patterns or home occupancy (Zheng et al. 2018). Such unawareness is carried to the privacy risks that non-A/V devices pose, thus underestimating the potential of inferring sensitive information from these devices. Users often follow habitual patterns in the use of devices—for example, turning lights on in the evening or using a certain kind of device at a certain hour. There are several risks associated with predictability since attackers can use these routines to determine when users are home or away, making them more vulnerable to physical break-ins or other malicious activities (Zou et al. 2023). Moreover, such risks can be even further worsened by social engineering methods; the attackers may pretend to be service providers or use phishing techniques to gather detailed information about users (ibid). These actions allow sophisticated attacks and increase the possibility of successfully compromising the user's security.

Another common issue is the ignorance of privacy settings, mostly engaging in insecure practices such as weak passwords or un-updated devices. Most users often ignore or overlook privacy settings on most devices, especially when simultaneously handling several devices and applications (Tawalbeh et al. 2020). Leaving the default settings allows hackers to exploit the feature of the device’s security, leading to potential data breaches. Another worse scenario is when users believe that a device from a well-known brand will be free and secure from any vulnerabilities or threats (Zheng et al. 2018). Many users also depend on manufacturers' privacy protection mechanisms in IoT devices. The misplaced trust might, on the contrary, lead them to a lack of being proactive about steps that ensure the safety of their data. Tawalbeh et al. (2020) state that many users also overlook the protection of their IoT devices, especially in terms of remote access protocols including, Wi-Fi, ZigBee, and Z-Wave. Without proper restrictions and a strong security mechanism, users become vulnerable to exploitation by hackers. This makes it easy for hackers to establish malicious connections through weakly protected remote access points, compromising device security and user data.

Furthermore, seeking convenience, users often share their personal data with IoT devices, such as smart assistants or connected home systems. These extensive sharing practices can significantly threaten users’ privacy, especially if it is not adequately protected or shared with third parties without the user's knowledge (Nguyen and Vu 2023; Tawalbeh et al. 2020). Prioritising convenience and connectedness with IoT devices can expose users to risks due to their acceptance of default settings, resulting in less protected data for the user, driven by prioritised ease of use (Zheng et al. 2018).

Studies show that the lack of awareness is the most significant privacy threat factor for users. For example, Zheng et al. (2018) state that users are unaware of how AI technology analyses and infers sensitive information from IoT devices and AI applications, leading to an unintentional breach of privacy. The general lack of knowledge highlights the importance of increased user awareness and stringent security methods to protect sensitive data in IoT ecosystems and AI applications. Moreover, users rarely understand how much they compromise their privacy and security when they input sensitive data into cloud servers to be processed using deep learning computations. In the process, they lose control of their data and risk violating their rights under regulations such as the GDPR. The lack of knowledge of the AI capability to create detailed profiles of users based on their data can further compromise user privacy (ibid). Therefore, raising user knowledge and ensuring clear transparency on data practices will be crucial in reducing these privacy threats.

## 2.2 Privacy in IoT and AI Environment

The rapid advancement in the IoT and AI sectors raises challenges in maintaining individuals’ privacy. AI systems need to be trained on datasets to satisfy users’ needs; the ease with which private data can be fed into AI systems raises concerns over how such data can potentially be misused. Compounding this problem is the unique challenge IoT devices pose in stealthily collecting data without users' explicit consent or awareness, thus magnifying the privacy issue.

From a technical point of view, the report by Danezis et al. 2015, from the European Network and Information Security Agency ENISA is analysing privacy issues and attack techniques, which raised significant concerns. Adversaries can break user anonymity through the identification of sensitive personal information. The primary attacks on privacy, as analysed by Zagi and Aziz (2020), are side-channel attacks, Man-in-the-Middle (MitM) and replay attacks. For example, a side-channel attack is the leak of information in cryptographic operations to infer either the master or sub-key, thus breaching privacy. A MitM attack intercepts data during transmission and alters it, thereby causing a direct breach of privacy. In a replay attack, the attacker retransmits or delays valid messages between two or more communicators, breaking into privacy and security.

Another serious threat Tao et al. (2019) identified is the collusion attack, where multiple insiders or entities may work together to bypass security defensive measures to access data and leak information. Eavesdropping is another attack in which attackers intercept unencrypted data during transmission to compromise privacy or even work with other entities, thus leading to several data leaks. Impersonation attacks are the attackers’ ability to pretend to be trusted users in order to access critical information. There is a risk of data leakage and destruction when data is intentionally or unintentionally exposed to unauthorised parties, which may result in the possible misuse of personally identifiable information. Several landmark security incidents—user privacy has been compromised due to data leaks with the implementation of IoT and AI. Privacy is not only a personal matter but also a security concern. It involves trust that other parties processing personal information will protect it well enough (Heurix et al. 2015).

Numerous incidents of privacy and data security violations have been reported in IoT and AI. For example, wearable devices—activity trackers—pose significant privacy risks due to sensitive data leakage and user re-identifiability. To comply with regulations such as GDPR, health data that these devices collect are typically de-identified and aggregated. However, reidentification risk is still present in most cases, especially when it is just partially aggregated (Pinchot and Cellante 2021).

In November 2017, the fitness app “Strava” released a global heatmap of athletic activity for users worldwide. Within months, these maps might reveal undisclosed military bases and other covert security operations, putting missions and soldiers’ lives at risk (Shastri et al. 2019). An analysis of the GPS data and movement patterns from one's Fitbit may be used to trace an individual's identity. This demonstrates risks in identification and privacy to be engaged in the use of personal activity data (Pinchot and Cellante 2021). Moreover, wearable devices’ location data or geolocations can further be analysed to monitor individual activities, raising privacy concerns (Pinchot and Cellante 2021; Zhang et al. 2020).

Zhang et al. (2020) state that wearable devices from Fitbit, Jawbone, and Google Glass collect sensitive information, such as sleep patterns, diet details, and pulse rates, which can put user privacy at high risk. Most of these devices have weak security measures, thus making them vulnerable to cyber-attacks and unauthorised access to users’ data. Additionally, their interaction through wearable device applications opens a range of vulnerabilities, such as Cross-Site Scripting attacks and SQL injection, making it more challenging to protect user privacy.

The growing use of connected wearable devices in the IoT has made user data, such as their habits, activities, and locations, vulnerable to being captured by third parties. There is also a risk that third-party application shared data will be linked back and re-identify users, which creates serious privacy concerns for personal information (Pinchot and Cellante 2021). The amount of information that can be leaked from users' daily activity through these IoT devices has become a significant privacy risk. Similarly, companies have slowly been moving away from the principle of data minimisation due to the extensive collection and storage of personal information. Research by Dwyer (2011) found that metadata could be used in the re-identification process on anonymised data or by correlating such data with other publicly available information. De Montjoye et al. (2013) conducted a study on mobile phone metadata, demonstrating that 95% of the individuals in a population of 1.5 million can be identified with only approximately 4 location and time data points. The risk of re-identification is magnified with IoT as attackers can easily associate anonymous IoT data with social media profiles or location data (Li and Palanisamy 2019). Most of these advantages of using the IoT rely heavily on collecting, storing, and processing massive quantities of data in real-time. Data aggregation from different sources also allows user re-identification, revealing personal identities and highlighting the failure of simple anonymisation techniques (ibid)

In terms of AI, Dilmaghani et al. (2019) show that sophisticated attacks, which include data poisoning, model inversion, and membership inference, have seriously threatened the information security and confidentiality of data used in AI. Data poisoning attacks involve the injection of malicious data at the training phase of AI models when AI algorithms make decisions based on sensitive attributes, which may result in biased predictions that compromise user privacy. This can cause false manipulated information generated by AI systems, leading to a lack of trust in AI applications. On the other hand, model inversion attacks allow attackers to make inferences on sensitive information based on analysing the output of models and thereby reveal, in many cases, private details about users. This attack highlights the AI model’s weaknesses to reverse engineering, emphasising the criticality of safeguarding training data. Membership inference attacks may potentially allow adversaries to determine if specific data points were part of the training set, possibly leading to revealing personal information. Additionally, through prompt injections and public text features, adversaries can use AI memory to extract personal data, which may raise privacy issues and legal actions against the technology. This issue underlines the importance of data privacy in ML, especially in protecting training datasets.

Integrating AI technologies and the need for massive data collection with ML has escalated significant privacy concerns, with hackers’ ability to exploit vulnerabilities to gain access to sensitive training data. For instance, Shokri and Shmatikov (2015) highlighted privacy risks in deep learning by proposing a method where multiple parties collaborate to train a neural network by sharing learning parameters while keeping their datasets private. Although this approach aims to protect data privacy, it raises concerns about the potential leakage of sensitive information through the shared model parameters. This example, along with earlier research such as Latanya Sweeney's re-identification of anonymized health records (Sweeney 2000), underscores the ongoing challenges of preserving privacy in AI and ML environments. This was accomplished by cross-referencing an anonymized medical dataset with publicly available information from the Cambridge electoral roll, which contained personal identifiers, enabling it to match and re-identify the supposedly anonymous health data. Another example is Fredrikson et al. (2015) discovered personal images from training data using a computer vision classifier that could reproduce almost 80% of an individual's image in the training set. Another example tested by Wang et al. (2023) shows that the GPT-2 model by ChatGPT has the ability to memorise sensitive data from the training set, potentially leaking information such as email addresses and private credit card details.

Moreover, advanced AI technologies raise significant risks to privacy, specifically biometric identification systems and smart home personal assistants, highlighting the need for strong safeguards of privacy and data handling practices. For example, the recent discovery of over three billion scraped facial images from social media networks by “Clearview AI” emphasises significant privacy risks related to unauthorised access and use of biometric data (Smith and Miller 2021). This technology combined these pictures with a facial recognition algorithm, greatly enhancing its functionality without adequate consent; therefore, it failed to comply with the CCPA and the Illinois Biometric Information Privacy Act (IBIPA). Similarly, Smart Home Personal Assistants, like “Amazon's Alexa”, have serious privacy worries due to continuous active voice channels and the complex architectures involving AI features. These continuous listening microphones may also result in the unintentional capture of recordings, as with Alexa, which had been recorded by accident and sent intimate conversations to an unauthorised recipient (Lau et al. 2018). The architecture of design in SPAs requires them to be continuously listening in order to respond to voice commands, making them prone to accidental activation and recording (Edu et al. 2020). The complex architecture and the integration into various technologies make them vulnerable to potential security breaches. The use of AI means processing large volumes of data that the individual is unable to control, leading to potential data exposure and misuse. Such threats at home could lead to serious privacy situations with actual impacts, such as identity theft, stalking, or unauthorised surveillance, enforcing the necessity of a solid methodology to safeguard personal data and users’ privacy (Lau et al. 2018).

## 2.3 Key Enabling PETs Technologies

PETs are important technologies and techniques to ensure personal data are secure. These technologies enable organisations to comply with privacy laws like the GDPR and CCP. Chaum (1981) was one of the first to use the term PETs in a research publication; it describes pseudonyms and anonymous approaches for unobservable communications over networks. These concepts are still in use today, such as anonymous communication and email systems (Cha et al. 2018). The work of Ziegeldorf et al. (2013) studies how current PETs can address different privacy threats in the IoT ecosystem; these include identification, localisation and tracking, profiling and linkage. The implementation of PETs in ML would secure the process by converting sensitive data into valuable information. This guarantees that the information is only disclosed to the concerned parties, keeping it secure from third-parties access (Ustundag Soykan et al. 2022). PETs lead to minimal collection of personal data, security during processing, and transparency in data handling practices (Cha et al. 2018). PETs consist of different technologies and techniques, for example, encryption techniques, which convert data into a coded format that is only decodable using a specific key; anonymisation, meaning removing all the identifiers from the data sets; differential privacy, which adds random noise preventing the identification of users; federated learning, where ML models are trained on distributed devices without exchanging their raw data with the central server; and secure multi-party computation, which allows joint computation over inputs. Such technologies will help organisations manage the risks related to personal data processing, protect against potential unauthorised access, and ensure compliance with privacy laws that result in trust and confidence gained from users.

### 2.3.1 Anonymisation

Data anonymisation is the process of removing personally identifiable information to keep individuals’ identities unknown. Privacy is protected by deleting identifiers, which could link an individual with stored data, protecting them from being re-identified. Anonymous data can be used effectively for analysis, research, and other purposes. This technique is highly required in areas where data privacy is critical, such as health, finance, and social data. Most privacy laws and regulations require anonymisation in order to protect personal information. Different advanced techniques in data anonymisation, such as K-anonymity, L-diversity, and T-closeness, provide strong privacy by mitigating several vulnerabilities. K-Anonymity ensures that one individual cannot be distinguished from K-1 others, removing the possibility of re-identification. L-Diversity guarantees that there is diversity in each group with respect to the sensitive attribute. T-Closeness attribute distribution similarity to the whole data set.

### 2.3.2 Differential privacy

Differential privacy (DP) is one of the widely used privacy-preserving paradigms that ensure inclusion or exclusion of any information related to an individual in the dataset has a minimal effect on the overall output (Zhu et al. 2021). In other words, it offers a solid mathematical guarantee that the outcome from a differentially private method will not reveal whether the data of some individual was used to generate the result. The concept of DP was introduced by Dwork in 2006; DP has become one of the fundamental approaches to ensure the privacy of individuals in statistical databases. DP has been applied in a wide range of areas. In the field of health, for instance, where confidentiality and privacy are critical to patient information, DP techniques allow researchers and data analysts to obtain aggregate statistics from databases without compromising the actual records of patients (Dwork and Roth 2014).

The basic idea of DP is to add carefully calibrated noise to the results of a query to avoid leaking sensitive information about individuals. This technique balances data utility and privacy protection, allowing for meaningful insights without breaching individual confidentiality. There are different noise addition mechanisms in data, such as the Laplace, Gaussian, and Exponential distributed mechanisms, which add random noise to hide the identity of individuals. Laplace and Gaussian techniques are meant for numerical outputs, while the Exponential mechanism is for non-numerical ones (Zhu et al. 2021; Dwork and Roth 2014).

DP is a method that adds noises during training while keeping individual records private. Xu et al. (2019a) propose EdgeSanitizer a deep-learning solution that uses local differential privacy (LDP) for mobile data analytics. This method guarantees the confidentiality of personal data while inferring useful information. EdgeSanitizer injects Laplacian noise into the learned process to protect sensitive information. It differs from the baseline approach by adding the Laplacian noise after the deep learning base, while the baseline approach adds the noise to the raw data. EdgeSanitizer protects each user’s training data under a single-user setting. EdgeSanitizer results in a high accuracy and utility reach of 87.27%.

Abadi et al. (2016) presented a deep-learning DP stochastic gradient descent (SGD), one of the first successful implementations of DP in ML. It involves adding random noise to the model updates during each iteration of training to ensure privacy. To ensure no single data point has a huge impact, the model limits the size of each update before adding noise, so that no update can overly influence the overall training process. Abadi and his colleagues introduce Privacy Accountant and Moments Accountant as one of the greatest concepts in DL-DP. They work together to track the aggravated loss of privacy during the training process to ensure the predefined privacy budget not exceed. These techniques have made it easier to train complex models while still protecting sensitive data, marking a significant step forward in privacy-preserving technology.

To reduce communication overhead and the amount of data transported to the cloud, Yang et al. (2018) employs a multifunctional aggregate framework that allows for various statistical aggregation functions, including both additive and non-additive types, using DP in the fog system. This method enhances the ML efficiency and ensures privacy. During the local aggregation process in the fog system, which is a cloud server close to the data source, noise is added to ensure privacy. As a result, the data are anonymised before transmission to the central cloud server. This process can reduce the transmission load since only the anonymised, aggregated results are sent to the central server. Efficiency and scalability are enhanced by distributing the computational load across multiple fog nodes making the algorithm practical for real-world applications.

### 2.3.3 Secure Computing

Secure computation is defined as a science of safeguarding communication and data with mathematical means, primarily to maintain data confidentiality, integrity, authentication, and non-repudiation. Homomorphic Encryption (HE) and Secure Multi-Party Computation (SMPC) are advanced cryptographic algorithms. HE allows computations over encrypted data without revealing it. SMPC allows many parties to jointly compute a function over their inputs without revealing those inputs to each other. HE and SMPC are two core cryptographic methods that increase the security and privacy of digital information.

#### 1) Homomorphic encryption

Homomorphic encryption (HE) is the ability to perform mathematical operations on encrypted data. The outcome of decrypting the data is identical to performing the same operations on the unencrypted data. This capability does not require access to decryption keys to process sensitive information, thus maintaining data privacy and security. There are three main types of HE schemes that differ in what capabilities they offer Partially Homomorphic Encryption (PHE), Somewhat Homomorphic Encryption (SHE), and Fully Homomorphic Encryption (FHE).

In a PHE scheme, only one operation, addition or multiplication, supports a particular function. For example, the (Rivest-Shamir-Adleman (RSA) scheme supports homomorphic multiplication (Rivest et al. 1977), while the Paillier cryptosystem supports homomorphic addition (Paillier 1999). SHE differs from PHE since both addition and multiplication can be performed on the ciphertext (Gentry 2009). Nevertheless, noise is a critical problem for SHE since each operation adds some noise to the ciphertext. This noise accumulation limits the number of operations that can be done making SHE not practical for any complex computation. Craig Gentry's work on SHE was an important step towards developing FHE. The FHE scheme allows an unlimited number of additions and multiplications on encrypted data, free from decryption. Bootstrapping is a technique that reduces the amount of noise within the ciphertexts (Gentry 2009). Gentry reduced the noise by enabling re-encryption of the data under a fresh key, enabling computations without the noise growing.

The IIoT is used in many industrial areas, including healthcare, smart manufacturing, smart city and smart grid. Li et al. (2021) designed a data privacy-preserving system for IIoT. The system uses lightweight HE to ensure the data is secure from insider threats and other data breaches, protecting the data even if there is unauthorised access. After the data is encrypted using a secret key, each piece of encrypted data is associated with a unique label, ensuring that only authorised users can have access. The IIoT have limited computing resources, which makes it hard for them to perform heavy calculations. For this reason, the encrypted data is uploaded to a cloud server. The proposed system offers a balance between keeping data private and being able to compute efficiently, where the server can perform computations on this encrypted data without decrypting it. The proposed system was tested using different sensors to calculate the air quality; the system does not learn anything about other sensors or anything about the data users. Overall, this research shows how data is managed securely and privately, ensuring that everyone’s information stays safe.  The currently existing HE schemes have been found to be incompatible with Wearable Devices, which are resource-constrained in terms of computing power, bandwidth, memory, and storage.

Huang and Duan (2024) proposed a privacy-preserving decision tree (PPDT) scheme in which multiple wearable devices participate; each of them contributes a portion of the data that needs to be evaluated by PPDT. These different devices work together to get more accurate evaluation results while keeping their individual data private. The use of transciphering technology and FHE optimises performance, making it suitable for devices with limited resources. The PPDT evaluation scheme with two cloud servers, one as a Trusted Third Party (TTP). The TTP provides every data owner (user) with a secret key to encrypt his health information. This encrypted data is further sent to the Computing Server for further processes. TTP generates the required keys and ensures that the data remains secure, therefore allowing the Computing Server to perform operations on encrypted data without ever getting access to the raw sensitive information.   
 

#### 2) Secure Multi-Party Computation

Secure Multi-Party Computation (SMPC) is a cryptographic method that enables the computation of a function over multiple parties' inputs without revealing those inputs to each other. It includes the process of computation division, where every party computes a small piece of the computation independently with their private data. After that, all these processed input party is combined without any data leak.

The first definition of the SMPC problem was introduced by Andrew Yao in 1982 in the "Millionaire's Problem," where he proposed a protocol for secure two-party computation (S2PC). This protocol allows two parties to compare their wealth without revealing their actual values, using the ideas of garbled circuits and oblivious transfer. This work has led to further research in this area as well as established the theoretical frame for secure multi-party computation among more than two parties (Chuan Zhao et al. 2019). With the advancement in cloud computing, mobile computing, and the IoT, the interest in SMPC has been renewed due to its feasibility in handling privacy and security issues in these sectors. SMPC has been applied in different domains, such as voting protocols, privacy-preserving ML, and collaborative data analysis.

SMPC relies heavily on network communications because it involves multiple parties, often spread across different locations. Each party holds a piece of private data that needs to be processed in several rounds. Each round involves a local computation followed by a communication step, which makes each party wait for others’ data to be processed for every round. The number of rounds and the complexity are the main factors limiting the SMPC performance. Bautista and Akkaya (2022) propose a pipelining-like approach for each round’s computation and communication by dividing the data, ensuring that the privacy of each mutually distrusted party is not compromised by the others. The pipelining-like approach allows the processing of multiple rounds simultaneals. The proposed system optimises the SMPC efficiency by using matrix multiplications, which divide the round data to be processed in sub-rounds. Matrix multiplication is a core component in Deep Learning (DL), that increases privacy-preserving computation.

In a Smart Grid (SG), the data of energy usage is transitioned through smart meters (SMs) to the utility company which can pose privacy risks to the consumers. Tonyali et al. (2018) propose a privacy-preserving SMPC protocol combined with FHE and a hierarchical aggregation approach that safeguards consumer data while keeping the data integrity without the need for a trusted third party. The proposed system discusses various methods and tools used to keep data secure while it is being processed and shared among different parties. The FHE allows computations to be performed on encrypted data, protecting it against unauthorised access and breaches; the data then is transmitted from the SM to local aggregators (such as neighborhood-level gateways). Allowing for the hierarchical aggregation process by collecting data at various levels within the network. This will reduce the volume of data transmitted across the network, optimising bandwidth usage and improving overall system performance. For data aggregation, Shamir's Secret Sharing (SSS) is used, which splits a secret into parts; only when a certain number of parts are combined can the original secret be revealed. To enhance the level of data security, the Pseudorandom Permutation protocol is used to shuffle data randomly, making it difficult to determine the original order. The SMPC protocol securely aggregates data from various meters without decrypting it, reducing the amount of data that needs to be transmitted to higher levels within the network. This ensures efficient data management while preserving privacy.

In a cooperative ML model, membership attacks pose a privacy threat. PRICURE is a system proposed by Jarin and Eshete (2021) that aims to protect the privacy of the model parameters and the input samples’ during collaborative inference among multiple model owners. The test results of PRICURE on different datasets show effective privacy protection and reduce the risk of inference attacks. SMPC is used to ensure that sensitive inputs remain private during computation, allowing parties to compute functions without revealing their inputs. The system splits the secret into random parts and shares them among multiple parties using additive secret sharing. This method ensures that a certain number of shares (t) are needed to reconstruct the original secret, enhancing security during computations. After computing the intermediate results using respective parts of the models and the sample. These results are sent to a central aggregator, which combines them into a final outcome. The final aggregation is protected against membership inference attacks, using DP, where the Laplace mechanism adds noise, enhancing privacy protection for both model owners and clients.

### 2.3.4 Federated Learning

ML is used in various fields that rely on centrally managed datasets for model training, posing serious privacy risks. The development of decentralised learning frameworks that allow collaborative model training without the exchange of individual datasets can handle these risks. For instance, Federated Learning (FL) is capable of scaling to thousands of participants while preserving data privacy.

FL is a decentralised ML approach where training data remains within the domain of each participant, thereby being ideally suitable for sensitive applications like healthcare and financial services. Instead of sharing raw data, participants train local models and share only model parameters. In this way, FL can be used in the IoT and smartphones to train a shared model locally in order to address the security risks associated with central data storage. These devices compute their updates locally and send them to a central server that aggregates them into a global model, maintaining user data privacy and security.

Konečný et al. (2016) proposed the theoretical foundation of FL as a possibility of training models in a decentralising method. After that, McMahan et al. (2016) proposed an efficient communication-learning algorithm called federated averaging (FedAvg) to average the updates of several clients, eventually reducing communication costs and handling non-independent and identically distributed (Non-IID) data.

An example of FL's practical application is using public Reddit data from “Google BigQuery” to train a next-word prediction recurrent neural network. This demonstrates FL's ability to handle real-world tasks while enhancing privacy by keeping data local (Konečný et al. 2017).

To prevent inference over messages exchanged during training and the final trained model. Truex et al. (2019) propose a hybrid approach that combines DP and SMC, an HE method called the Paillier cryptosystem, to enhance privacy in FL. By implementing this technique, the proposed system reduces noise injection as the number of parties increases. The amount of noise added by each participant depends on how much privacy is needed and the sensitivity of the query. Then the data is encrypted using the Paillier cryptosystem, which allows securely aggregating the responses.  This system guarantees pravicy, as the number of participants increases, the system reduces the noise injection. However, Using the DP method in FL can make the model less effective since each party will train his version of the model. Most of the methods used in FL depend on trusted aggregators where the central party collects and combines the model updates.

Xu et al. (2019b) built a HybridAlpha that does not depend on a trusted aggregator, making it highly secure. HybridAlpha uses the SMC protocols that enable participants to compute their collaborative results, dropping the need for trusted aggregators. Using this encryption function can ensure that even if participants drop, the privacy of other participants is still guaranteed. HybridAlpha shows promising results by reducing the training time by 68% and data transfer volume by 92%.

Singh et al. (2021) proposes the use of Blockchain-based IoT cloud platforms and FL as Privacy-Preserving in Smart Healthcare while ensuring security and privacy. To preserve privacy, a noise will be added using DP and the models will securely aggregates using HE. All the encrypted model updates are sent to a central server, which will log them on a blockchain. The blockchain guarantees that all of the updates are immutably and securely logged, ensuring transparency in the record of all operations.

## 2.4 PETs Challenges

### 2.4.1 Challenges in DP

Balancing data utility and privacy is challenging when implementing DP in an IOT or AI system. The protection of data in DP is about the addition of noise, that may result in reducing the data utility (Zhu et al. 2021). For this reason, it is computing intensive and requires extensive resources. This calls for careful optimisation, balancing the tuning of epsilon (ε) and delta (δ) parameters. Epsilon (ε) is tuned to ensure a trade-off between the release's privacy level and utility, where lower values of epsilon offer more privacy but result in less utility, and vice versa. Delta (δ) must be adjusted according to the sensitivity of the data, so privacy is preserved in a collection of datasets and classes of queries.

Furthermore, applying DP within ML requires designing new algorithms that effectively trade off privacy with model performance. Integrating DP with other algorithms such as Secure Multiparty Computation in ML, may raise security issues (Dwork and Roth 2014). Real-world implementation of DP may be challenging due to the nature of data types (Zhu et al. 2021), such as location and trajectory, which are tracking data that can be easily distorted or lost. In addition, an epsilon (ε) parameter is hard to set and needs qualified human resources and proper computing environments to ensure the correct implementation of DP.

### 2.4.3 Challenges in Secure Computation

**1) Challenges in HE**

There are several challenges in HE implementations. The major difficulty is that FHE and SHE are mathematically complex and have not been efficiently implemented for practical use. The main issue lies in the fact that operations on encrypted data without decrypting are extensive and require highly complex computations (Alaya et al. 2020). Generally, the HE algorithms are slow when compared to traditional encryption techniques. Further, HE consumes a significant amount of memory in the process of encryption and decryption and during homomorphic operations, the processes involve large memory requirements; this might be a bottleneck while running in practical applications.

Moreover, the key management needs more attention; the protection of the keys generation and handling is important, especially in the cloud environment where data are stored off-site. The security of a system may be affected by exposure or loss of keys (ibid). Additionally, the process of securely transmitting and handling these large keys across numerous locations may lead to further risks and logistical problems. Moreover, large keys are part of the complexity because it is hard to manage and store them securely due to their massive size. Furthermore, the ciphertext size scales to be extremely large compared to the original plaintext. This highlights the enormous storage and communication overhead in data transmission in HE algorithms. The large memory requirements are challenging, especially in resource-constrained environments like mobile devices and IoT sensors.

**2) Challenges in SMPC**

SMPC has some challenges that should be resolved to make this technique practical and scalable in real-world applications. SMPC methods, such as garbled circuits (GC) and oblivious transfer (OT), have intense computational resource requirements (Feng et al. 2023). When the data volume is massive, time, memory, and communication costs rise exponentially, which makes it impractical to handle large-scale data. For example, the use of GC and OT comes with significant costs when the data size increases. Adding new pieces of data means a rapid increase in computational and memory resources required to process that data.

From a scalability perspective, such exponential resource consumption is excessive when training ML models with large datasets using SMPC approaches. Many SMPC protocols are computationally intensive, making it difficult for participants with less processing power and memory to contribute. Moreover, devices in mobile or IoT environments might not have the performance needed for these cryptographic computations to be implemented efficiently (Chuan Zhao et al. 2019).

The communication overhead in SMPC introduces additional latency, which is critical in real-time applications. Some studies have found that removing the communication cost of the inactive conditional branches helps reduce the latency (Feng et al. 2023). Communication latency is a huge performance bottleneck for SMPC protocols in geographically distributed settings. Some Real-time applications may need more than one interaction round to complete SMPC protocols, such as online auctions or collaborative data analysis.

Furthermore, SMPC protocols face significant challenges to support dynamic participation and preserving security. Another challenge for SMPC protocols is network failures or participant dropouts (ibid). In an ideal world, if a participant in the protocol is unresponsive during computation, the protocol should still be able to be completed successfully without compromising the security of the remaining parties.

### 2.4.4 Challenges in FL

FL is one of the promising methods used in decentralised ML; however, it may face several critical challenges. FL models use data from different clients, which highlights a prime issue about handling non-independent and identically distributed (non-IID) data. This may complicate the aggregation process and affect the performance of the global model (Konečný et al. 2017).

The interactions between clients and the central server have a significant impact on communication that tends to be costly and slow. High communication costs accrue especially in large networks with limited bandwidth, calling for effective communication strategies (McMahan et al. 2016). Moreover, FL is still vulnerable to attacks like model poisoning, where clients may send damaged updates to corrupt the global model (Konečný et al. 2016). Furthermore, FL clients—such as smartphones and IoT devices—have limited computational power, storage, and battery life, which makes intensive local model training challenging (Hard et al. 2019).

## 2.5 Privacy Challenges in IoT and AI Environment

### 2.5.1 IoT Challenges

IoT devices generate a massive amount of data, some of which contain sensitive personal information. Data may come in the form of context data, continuous data, or media data. Since IoT is a connected network of systems and devices, and these devices have limited resources, ensuring data privacy will be a challenge. Therefore, the main challenges of privacy in IoT are:

* **Large-Scale Data Collection**: IoT devices can collect vast amounts of personal and sensitive data, such as geolocation, daily routines, and health-related data. Users may be unaware of the information being collected and its consequences (Princi and Krämer 2020; Morel and Fischer-Hübner 2023).
* **Data Exchange and Connectivity**: When two different IoT systems communicate with each other, the information shared may pose the possibility of a data breach. Moreover, fitness trackers collect data and share it with third-party applications (Pinchot and Cellante 2021). Data transmitted to third-party services may lack a secure communication protocol or poor data handling.
* **Insufficient Safety Protocols**: Most IoT devices and cloud systems have inadequate security protections, which can even lead to possible hacking and unauthorised access. Such weaknesses in security might lead to the leakage of data through complex attacks, such as collusion attacks, where multiple malicious nodes join to compromise the network and damage security mechanisms. Additionally, data is prone to be intercepted during transmission between servers, systems, and devices. This unauthorised interception is transmitted to an external destination (Tao et al. 2019). Data breaches can further rise because of the connected nature of IoT, where if one device has been breached, it means breaching the whole system (Du et al. 2018).
* **Heterogeneous Devices and Networking**: The nature of the IoT ecosystem tends to be complex, considering the use of various techniques of communication, either wired or wireless technologies, including Bluetooth, Zigbee and LPWAN. This high mix of mechanisms of communications demands diversified security. Moreover, an IoT infrastructure has a large number of devices, ranging from sensors and switches to gateways and actuators. These differences need different algorithms and communication protocols for effective management of their operations (Zagi and Aziz 2020).
* **Ownership and Use of Data**: Aggregation and sharing of data are privacy and security issues for the IoT. For example, ownership rights over the data generated by users through their IoT devices may not be entirely clear. Information collected for one purpose can find its use in another without letting the user know about it (Nguyen and Vu 2023). This may violate the privacy of users and breach their trust since they are unaware whether their personal information is being used in the ways that they specify or permit.
* **User Control and Consent**: One of the most significant difficulties in implementing IoT privacy and security practices is users' lack of control over their data. It is challenging to obtain active, explicit, and informed consent from users regarding the collection and distribution of data. This leads to a transparency deficit that decreases trust and increases privacy risks for users (Morel and Fischer-Hübner 2023).

### 2.5.2 AI Challenges

Training AI models requires extensive data collection and complex decision processes, introducing several privacy risks. The major privacy issues related to AI can include:

* **Data Dependability and Quality:** AI systems and deep learning models rely heavily on vast amounts of training data, often containing sensitive personal data, which raises privacy risks (Bae et al. 2018). Biased or low-quality datasets can result in biased and low-quality AI outcomes. For example, in healthcare, such outcomes may lead to discrimination and raise ethical and privacy concerns. High-performing models must preserve data privacy by detecting and eliminating poor-quality data.
* **Minimising Data and Obtaining Informed Consent:** Due to the technical nature of AI, obtaining informed consent for user data can be challenging. People may not fully understand how their data is used when AI systems are involved. This lack of transparency may lead to informed consent issues, such as in healthcare AI systems where patients are unaware of how much their data is processed or shared (Gerke et al. 2020). AI systems often collect more data than necessary, raising privacy concerns. For example, algorithms can quickly repurpose data without renewed informed consent, leading to significant privacy risks, such as Facebook analysing user data to predict personal attributes also the leaked location data showing the movements of millions of phones in the US (Andreotta et al. 2021). These issues highlight the need for more transparent, more understandable consent processes in AI data collection.
* **Algorithmic Accountability and Transparency:** The AI algorithms make it difficult to understand the decision-making process. Transparency is crucial for building trust as it allows users to understand how AI systems make decisions. For instance, a lack of transparency when AI evaluates teacher performance based on student grades can lead to significant consequences, such as job loss (Andreotta et al. 2021). Ethical concerns related to bias and fairness are better addressed when decision-making processes are transparent.
* **Data Anonymisation and Re-identification:** Ensuring proper anonymisation of data is challenging, as there is always a risk of re-identification, where anonymised data can be linked back to individuals. This is a significant concern in data privacy, particularly with the rise of big data and AI technologies. Re-identification is possible, especially when the data is only partially aggregated. (Pinchot and Cellante, 2021). This risk raises ethical concerns, as re-identification can lead to discrimination (Andreotta et al. 2021). It also creates challenges for compliance with regulations like GDPR, as organisations may struggle to ensure that anonymised data remains protected.

## 2.6 Privacy and GDPR

The rapid growth of IoT and AI technologies has led to significant advancements but it has also raised substantial concerns for individual privacy. Many IoT devices in daily life function as passive data gatherers, which exacerbates worries about the intensive collection of personal data (Ghayyur et al. 2020). To address these challenges, strict privacy laws such as the GDPR and CCPA have been implemented.

Privacy preservation requirements, as outlined by Pfitzmann and Hansen (2010), include:

* **Anonymity**: This property ensures that a user's identity remains concealed from third parties, making the user unidentifiable within a group known as an "anonymity set."
* **Data Minimization**: Service providers should only collect and process the information necessary to provide a service. Collecting as little data as possible reduces the risk of user profiling and tracking. Data minimisation helps fulfil GDPR Article 5, which emphasises clearly defining the purpose of data collection.
* **Unlinkability**: The data collected during credential issuance cannot be linked to the user's identity during subsequent authentication. Multi-show unlinkability ensures that even if the same credential is used in different sessions, the service provider cannot correlate these sessions to track the user's activities or identity.
* **Unobservability**: This ensures that users' actions are protected from being observed or tracked by unrelated parties. Unobservability keeps user activities undetectable and anonymous, preventing third parties from identifying operations.

Several key standards, as listed by GDPR (GDPR.EU 2023), are important to follow in IoT and AI implementations:

* **Processing of Personal Data (GDPR Article 5):** This article forms the principles for processing personal data. First, it dictates that the purpose of collecting the data should be clearly and transparently defined. Personal data should be collected for specified, explicit, and legitimate purposes and not further processed in a way that is incompatible with those purposes. Secondly, the data should be kept in a form that allows the identification of data subjects for no longer than is necessary for processing the data, hence helping minimise the risks of data breaches and misuse of information. Moreover, data must be processed securely to avoid unauthorised or unlawful processing and accidental loss or damage. This enables transparency and limits the use of data to ensure it is not misused.
* **Consent Conditions (GDPR Article 7):** Under this article, obtain explicit consent from users. The notification of collecting personal information must be clearly defined with the option of either giving or refusing consent. It should be clear to the subjects who collect what type of data and the purpose for the collection, providing the user with control and transparency.
* **Records on Processing Activities (GDPR Article 30):** This article requires controllers and processors to maintain records of processing activities that are complete and descriptive in nature. Such records typically contain, the purposes for which data was being processed, categories of data subjects and personal data, and details of data transfers. These logs are important for proof of compliance and accountability.
* **Security of Processing (GDPR Article 32):** According to this article, the processor must implement strong security measures for personal data processing by putting proper risk management procedures. Organizations should consider and manage the risks associated with the data that they hold and put in place adequate technical and organizational measures that assure appropriate protection against unauthorised or unlawful processing and accidental loss, destruction, or damage of personal data.

This includes minimising data collection, regular privacy audits, impact assessments, and transparency of data practices to assure compliance with regulations and earn user trust. PETs have become a major tool in enforcing GDPR-defined security requirements, helping provide the right balance in handling personal data processing risks through effective measures that lead to robust solutions against unauthorised access, data breaches, and other security threats. Privacy-preserving methodologies that enable the benefits of IoT and AI without compromising users' privacy should be widely available. Such methods are based on cryptographic techniques and ML approaches. Deployment of privacy-enhanced technologies, including federated learning for training AI models across decentralised devices without data transfer, and introduction of differential privacy by adding "noise" to data, could further be used for the protection of personal information. Homomorphic encryption and secure multi-party computation are cryptographic techniques that enable computations on encrypted data in such a way that does not reveal sensitive information.

Enterprises can effectively and responsibly use AI and IoT in their services while ensuring the users' privacy is taken into account and that trust is fostered through the use of privacy-preserving techniques. This balance is key to an organisation where trust and regulatory adherence have become essential (Ghayyur et al. 2020).

# Chapter 3: Research Methodology and Implementation

This chapter details the research methodology employed in designing and implementing an adaptive encryption scheme for IoT sensor networks the primary objective of this research is to develop a method that classifies network traffic data into criticality levels and applies appropriate encryption techniques based on the classification of the methodology is structured into several stages including data collection feature selection model training encryption scheme development and performance validation.

## 3.1 Research Design

The research methodology employed in this study is a blend of quantitative and qualitative approaches leveraging open-source datasets to design develop and evaluate an adaptive encryption scheme for IoT sensor networks, this chapter details the systematic procedures followed including data processing model development and encryption implementation the methodology is designed to ensure the validity and reliability of the findings providing a robust framework for the analysis.

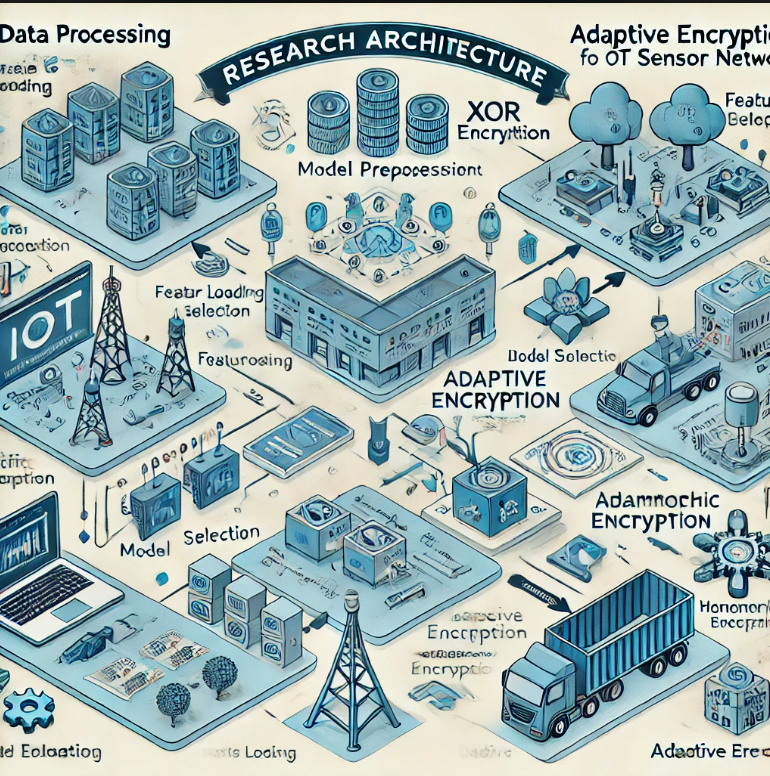


Figure 1 Research Architecture for Adaptive Encryption Scheme

The research design is divided into three main components:

1. **Data Processing**, which includes data acquisition and preprocessing
2. **Model Development**, where a Logistic Regression model is trained to classify data
3. **Adaptive Encryption**, which applies XOR or homomorphic-like encryption based on the data’s classification.

## 3.2 Data Sources and Collection

The data set used for the current work is the network traffic data which containing both the normal and the attacks traffic record. The dataset was collected from the N-BaIoT dataset which is public and contains several features that describe the behaviours of the network. The data was split into two categories: benign traffic with the normal network activities and malicious traffic with the security threats or violation. All the data points in the dataset are labeled based on the type of feature and hence it supports supervised learning.

To process the data, the two sets of benign and malicious data were merged into one set with a binary label of criticality level as 0 for benign and 1 for malicious. The input dataset features were preprocessed and after a detailed feature reduction stage, six features were considered to be relevant in flagging between benign and malicious traffic. After that, each of the features was scaled using a StandardScaler in order to normalize the training data for the model.

## 3.3 Quantitative Methods

The quantitative part of the study proposes and tests the creation of a Logistic Regression model in order to increase the accuracy of classification of network traffic into low-critical and high-critical traffic. The following steps outline the quantitative procedures:The following steps outline the quantitative procedures:

* Data Processing: The features were selected and normalized in this phase so as to enhance the model training process from the dataset. The features where therefore chosen because of their applicability to traffic analysis in various computer networks.
* Model Development: After the pre-processing of the data, the Logistic Regression was used to learn on the data in order to classify the data points. The regular assessment tools and parameters for measuring the best performance of the item were employed which included accuracy, confusion matrix, and ROC-AUC score for the model. The quantitative findings gave a promise on how the model may identify various degrees of data criticality.
* Performance Evaluation: The accuracy and the misclassification of events in the model were described in detail. Moreover, the ROC curve was used to assess TP and FP co-efficients and AUC score which gives the single quantitative evaluation of the results.

## 3. 4 Qualitative Methods

The qualitative part of the research is based on the interpretative approach to the examined encryption strategies for the classified data. The qualitative methods included:

* Encryption Strategy Analysis: There were two types of encryption mechanisms used in the study: Am XOR encryption for low critical data and a homomorphic encryption like method for high critical data. These methods have been chosen depending on their suitability and efficiency in protecting IoT network traffic. These encryption processes were then reviewed with the view of identifying their advantages and disadvantages especially in environment where devices are constrained in terms of the resources they hold such as IoT devices.
* Process Validation: For the enhancement of the encryption scheme reliability, the qualitative validation of the encryption results was conducted in the research. To do this, the research evaluated if the applied encryption methods matched the obtained criticality levels as predicted by the model in its encrypted outputs.

To further understand the model's performance, several visualizations were generated:Besides that, the following graphs were produced to describe the details of the model’s performance:

1. Confusion Matrix: The confusion matrix was also quite useful because it gave an elaborate report of any model including the true positive, true negative, false positive as well as the false negative. This matrix was inevitable in as much as the work was able to position the model in such a hostile environment where the low critical data and high critical data were well segregated; the recorded false signals; the false positives and a fairly acceptable window of false negatives.
2. ROC Curve and AUC: In order that the trade off between the true positive rate and the false positive rate could be observed, a ROC curve was plotted. More importantly it was observed that the use of yield AUC converged concentration of vs curve by equal to 1. 00 which indicate that the model is placed in a position to maximize on the separation of the two classes. That sort of AUC score portrayed the findings the accuracy of the full sample datum’s in parasitic decreased and other correspondent datum’s were estimated very surely for identification.
3. Feature Importance: Certain quantification of the features of cohorts for the significance of the logistic regression model was made referring to the coefficients of features used in the model. This analysis gave an understanding of which of the features had a higher given impact in determining the criticality of data as the subsequent iterations of the model and selection of the right features was being understood.

## 3.5 Research Implementation

The implementation phase of this research involved a comprehensive approach to integrating the developed adaptive encryption scheme within the IoT sensor network environment this section details the various stages of implementation including dataset preparation model training adaptive encryption and validation processes.

### 3.5.1 Dataset Preparation

This process started with the pre-processing of the dataset which was an important step in the entire process of Logistic Regression model implementation and the cryptographing of the information. The raw data, that is, benign and malicious network traffic data, was then implemented in the environment. They were used to generate low critical (benign) and high critical (malicious data labels). Key features relevant to network traffic analysis were selected for model training, including.

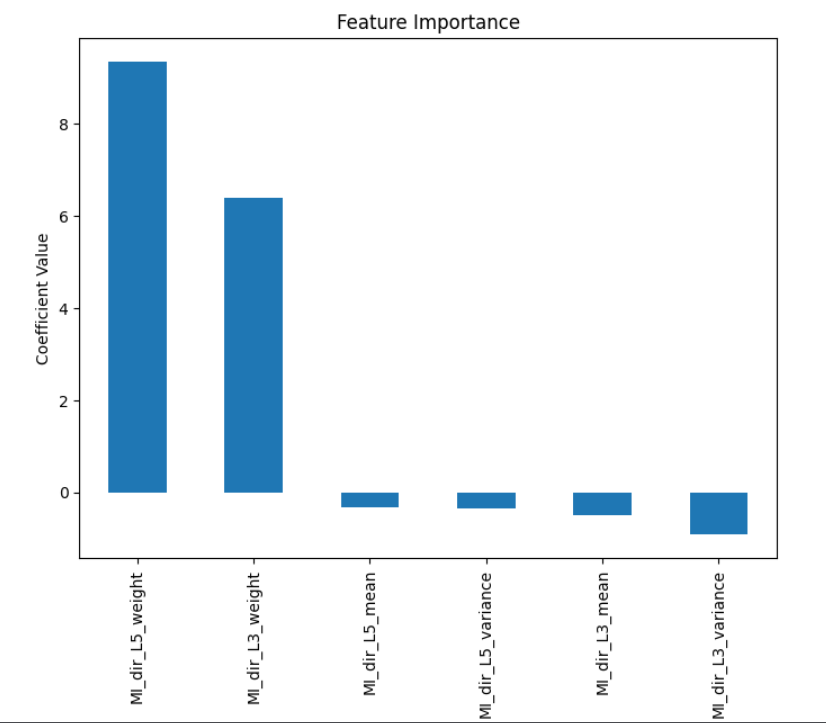


Figure 2 Features importance

In figure 2, the importance of each feature was assessed by examining the model s coefficients this analysis provided insights into which features were most influential in predicting data criticality allowing for further refinement of the model and feature selection process in future iterations

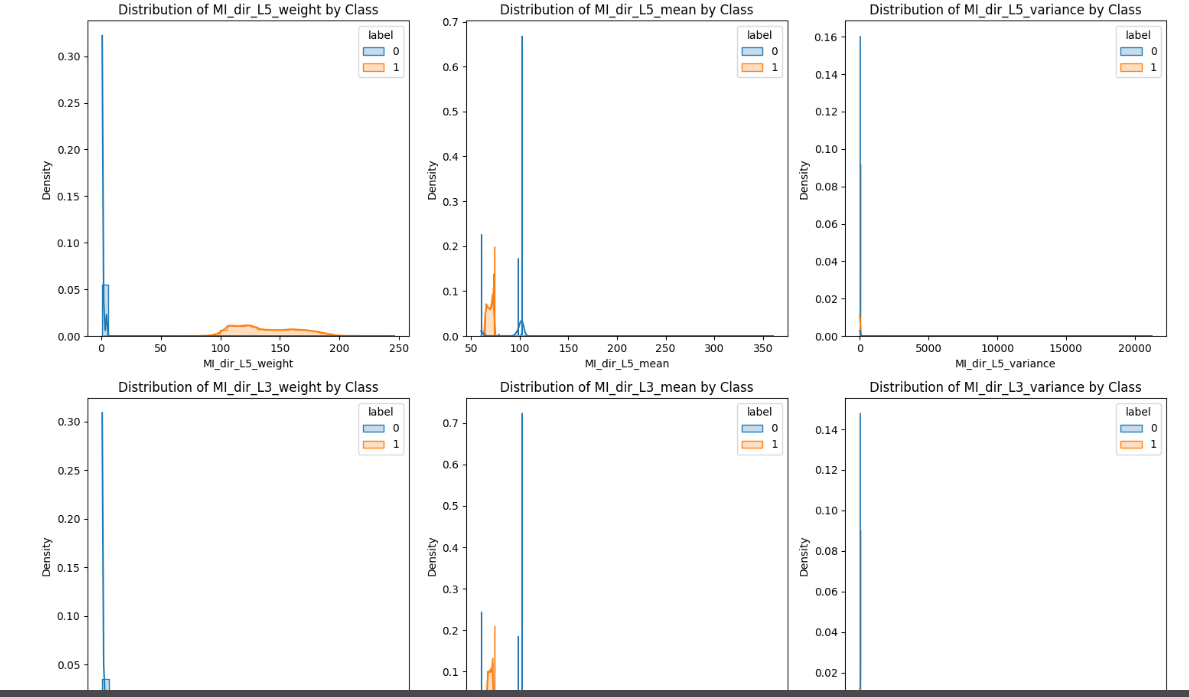


Figure 3 Data visualization before training and testing

These features were selected based on the level of impact that they imposed on the accuracy of the model.

### 3.5.2 Model Training

For classification of the network traffic data to be normal and or abnormal, in this work, we employed the Logistic Regression model. First, the data was divided into a training sample and a testing sample: First it had the former category, which comprised of 80% of the cases, then the latter had the 20% balance. Here this split was made in such a way that the model could be tested with data that it was never exposed to in the training in a bid to enhance the virtuosity of the former. The features were normalized by centering and scaling them, in order to help during model convergence and in feature selection during the training of the model. Standardised training data use the Logistic Regression model with specified random seed to make the results reproducible during the training of the model. Then the accuracy on every test set was measured with sensational accuracy rate at the end of which it was 99%. This high accuracy means that the model can isolate the benign from the malicious traffic, meaning it is effective when used for this purpose.



Figure 4 ML model output

### 3.5.3 Adaptive Encryption

The centerpiece of the implementation was adaptive encryption scheme, conceived to use different encryption algorithms on the basis of the classified criticality of the data flowing in the network stream. Due to these reasons, this is the most flexible approach that must be used in IoT since the choice of the encryption method relies on how sensitive data is and the computational capacity of the device.

#### XOR Encryption for Low-Critical Data:

For data classified as low-critical, a lighter XOR kind of encryption technique has been used. This method simple though effective in its way, is only applicable to low-critical data since it does not involve any computationally intensive process. This was done by means of a predefined key for the XOR encryption function to perform operations on the data required for encryption and decryption without imposing considerable load on the resources of IoT devices.

#### Homomorphic Encryption for Like Method for High-Critical Data:

As for high-critical data, the method we used somewhat resembled homomorphic encryption. Although not a fully homomorphic encryption, this method simulated the process by appending a unique encrypted string based on the data's hash value. This approach was adopted to point to the fact that high levels of security are being exercised when dealing with the data such that even if the data is intercepted the information is secure.

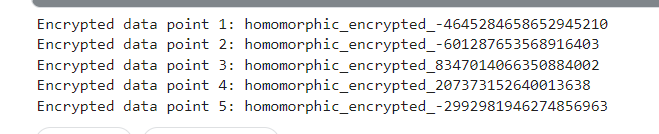


Figure 5 Adaptive Encryption output

In another experiment to solve this problem and for coping with the processing load especially if the GPU is not an option, the encryption was done in batches. This batch processing enabled the encryption scheme to handle large data quantities because the quantities were addressed in batches, and not all at the same time. The encryption simulation for high-critical data is working as expected, producing unique "homomorphic encrypted" strings based on the data's hash values. These strings represent the encrypted form of data points, simulating the behavior of a real homomorphic encryption scheme.

### 3.5.4 Validation Process

This was succeeded by an adaptation of the encryption that was followed by a verification step to gauge the outcome of the classification, as well as the suitability of the adopted encryption techniques. By so doing, the validation, in this case, demands that one should set a portion of the data aside and see whether the classification by the Logistic Regression model corresponds to the actual classification.

The validation also provided the information on which category the data point belonged to as well as the kind of encryption the data point was assigned by the adaptive encryption system indicating on the manner of application of the adaptive encryption. The validation process revealed a mix of correct and incorrect classifications, with the following result. The validation of the results showed that there were both right and wrong cases classified as such with the following outcomes:

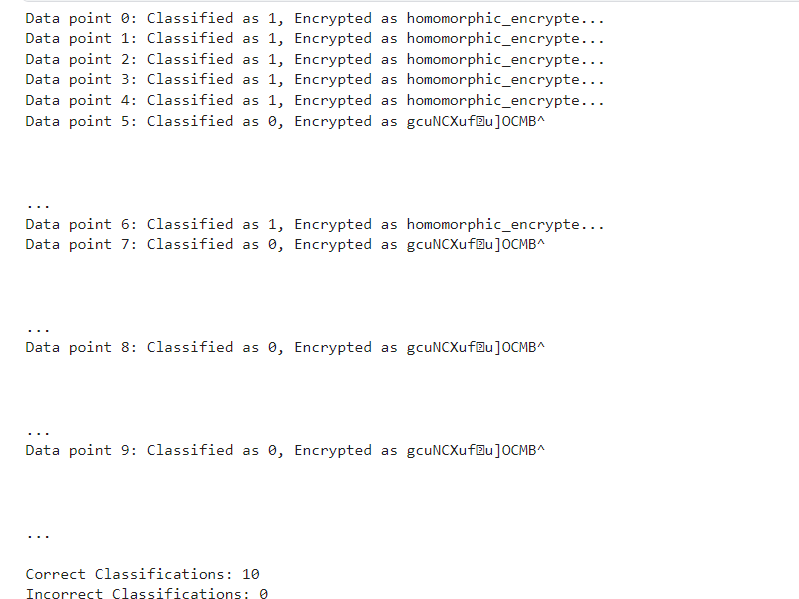


Figure 6 validation output

These effects demonstrated that specific aspects of the adaptive encryption scheme could be enhanced so as to augment the effect of minimised misclassification through the model.

## 3.6 Conclusion

The implementation of the adaptive encryption scheme within the IoT sensor network environment was a multifaceted process that integrated quantitative and qualitative analysis methods from dataset preparation to model training adaptive encryption and validation each step was meticulously planned and executed the resulting encryption scheme while effectively also highlighted areas for further research and improvement particularly in enhancing model accuracy and exploring more advanced encryption techniques for high critical data this detailed implementation serves as a foundation for future research in secure IoT communications where adaptive encryption schemes can be further refined and adapted to meet the evolving security needs of IoT networks.

# Chapter 4 Results and Discussion

In the Chapter 4 “Results and Discussion”, we present the evaluation and validation of the proposed adaptive encryption for the IoT sensor networks. This section describes the results of the encryption which has been applied, the results measure the performance of the encryption process against other techniques to check its feasibility and effectiveness.

## 4.1 Model Performance Evaluation

Using the findings of the study, the Logistic Regression model used in the process of differentiating network traffic data to either normal or abnormal, high performance was noted in the test phase. A divide was then made with the data set by using the training and testing (at the ratio of 80:20) On the training data, all the features were normalized which meant that all the features were rendered to zero mean and unit standard deviation. This feature standardization was necessary to reduce the case of standardisation where some of the respective model’s features could overpower the entire output due to scaling differences.

Once the entire study was complete and the test set passed through the model, the accuracy was pegged at a stunning 99 percent level. 97%. As far as this result is concerned it could be analyzed that nearly all types of the network traffic were properly identified by the model and were differentiated between normal and abnormal. This high accuracy the extent to which Logistic Regression is capable of performing this classification since the model is scalable in regards to features preprocessing/ scaling.

To further understand the model's performance, several visualizations were generated. However, to have a detailed and a clear picture of how the model is performing, several other visualizations were created:

### 4.1.1 Confusion Matrix for results analysis

The confusion matrix analysis of the performance of the model goes on to support this having observed that the model exhibited high accuracy in the differentiation of the High-Critical and Low-Critical network traffic data. In particular, the model obtained the number of TP is 24,563 reflecting the High-Critical data points and the number of TN is 9,852 reflecting the Low-Critical data points. More to the point, the model yielded an FP rate of zero, which means that no Low-Critical data points were classified under the High-Critical label, a factor which is of profound importance when dealing with actual applications. Similarly the False Negative (FN) count was equally low with the value of 10 hence implying that almost all High-Critical data points did not get to be underestimated or classified under Low-Critical. Such a level of accuracy underlines the versatility of the proposed model as its main objective is to reduce such serious mistakes where such applications as cybersecurity are dominant. In general, the outcomes from the confusion matrix support the fact that the present model is consistent and effective in its capability of isolate the various kinds of network traffics without much misclassification.

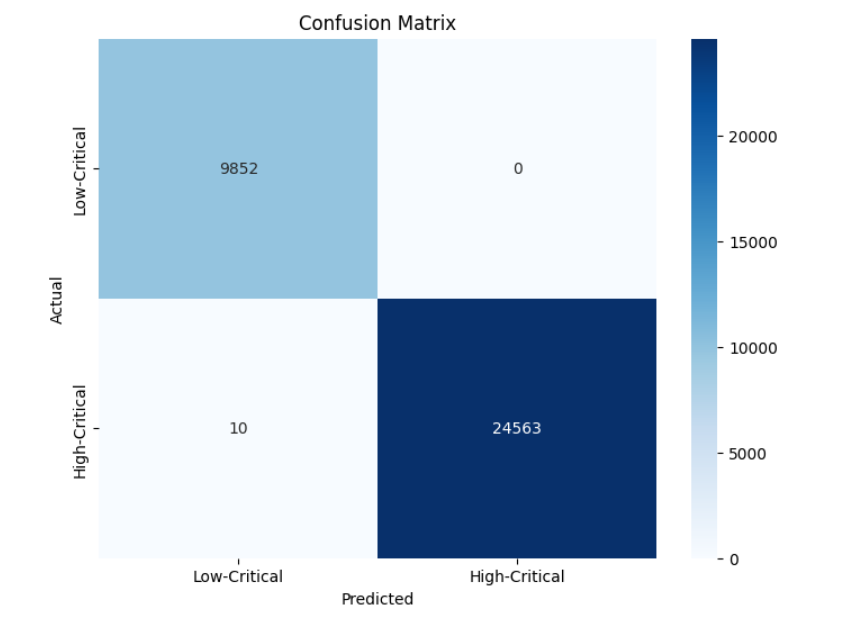


Figure 7 Confusion matrix

### 4.1.2 ROC Curve and AUC Analysis

The alternative diagnostic test is the Receiver Operating Characteristic (ROC) curve that is employed in the assessment of the performance of the model through a graphic representation of the True Positive Rate (TPR) against the False Positive Rate (FPR). In this analysis, the ROC curve is seen to be a very tight curve close to the top left corner of the graph; which is very desirable for any model. This positioning close to the top left quadrant means that the model’s performance is greatly optimized in being able to give high TPR without affecting the FPR sharply. In support to this, the AUC is evaluated to be 1. 00 being the maximum score that a student could obtain in the test. An AUC of 1. Assuming, that matrix 00 unequivocally testifies that the model is capable of giving an absolute classification between High-Critical and Low-Critical classes; the reader can conclude that the model exhibits an extraordinary performance in the context of the classification tasks. This kind of result indicated that the proposed model has achieved good performance and it became more applicable for essential application where the classification is so crucial.

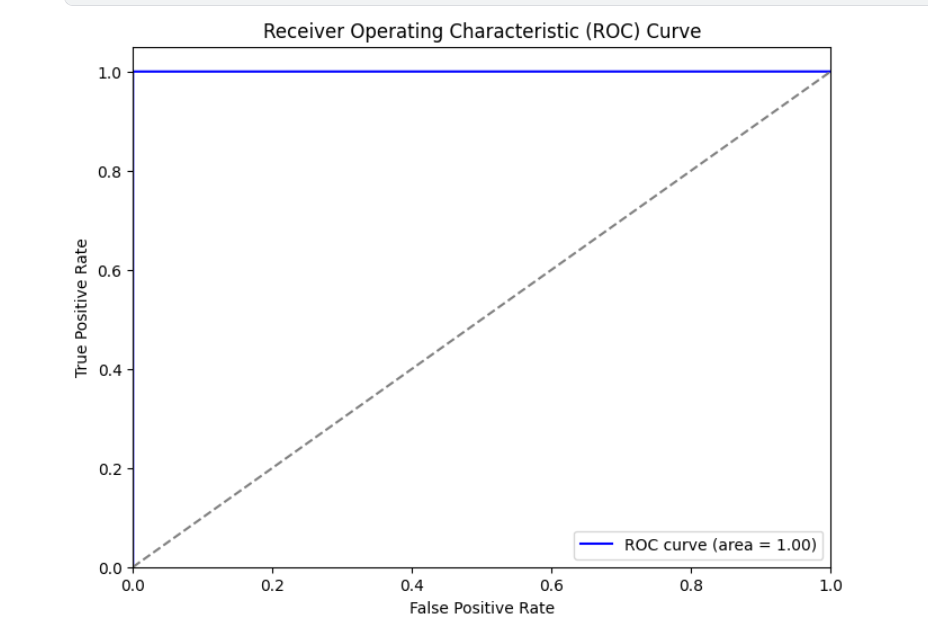


Figure 8 ROC curve analysis

## 4.2 Discussion of Feature Importance

The features considered for this study are MI\_dir\_L5\_weight, MI\_dir\_L5\_mean, MI\_dir\_L5\_variance, MI\_dir\_L3\_weight, MI\_dir\_L3\_mean and MI\_dir\_L3\_variance and all these features were selected bearing in mind the problem domain. Every feature becomes an independent characteristic of the network traffic and brings something individual to the decision-making process of the model. The high accuracy levels attained points to the fact that these features where indeed useful in capturing the hidden patterns that exist between benign and malicious traffic.

Supplementary to that, the density plots featuring each feature also separated by class labels gave additional information on the distribution of benign and malicious traffic. It was noted, for instance, that the two classes had similarities as well as clear differences in terms of several features. This distribution difference means while using Logistic Regression increasing coefficients of different predictors, it might impact learning linear relations in the data successfully, some nonlinear trends might not have been appreciated during the analysis and may be with more complex models in subsequent studies could be realized and exploited.

## 4.3 Generalization Of Their Model And The Likelihood Of Over-Fitting

The level of accuracy attained could be misleading, however, there is always the problem of overfitting. A model that gives very high accuracy on the training and test sets might have in fact ‘overfit’ the data, which means it has learned data specific patterns to understand the given data. However, due to the characteristics of the dataset and the fact that Logistic Regression is a linear model, risk of overfitting is somewhat low. The fact that the features of the model are standardized and that Logistic Regression inherently has more regularized characteristics than Decision Tree through the process of tuning parameters minimizes overfitting. However, even more accurate validation of the model in terms of generalization to new, not used data during the training or testing, is possible.

## 4.4 Comparison of the Model with Other Models

Indeed, it is important to compare the performance of the Logistic Regression model with other more complex models like Neural Networks or Random Forests or Gradient boosting Machines, etc. All of these models could presciently depict other higher order and curvilinear relationship in the data than what is exploited in Logistic Regression. Nonetheless, Outlook remains the simplest and easily interpretable model among all presented; its benefits are most apparent when the interpretableness or the faster calculations are vital.

## 4.5 Recommendations and Implementation

Below are the implications of the results obtained in this study with regards to real-world application of the work done and specifically in the area of Cybersecurity. To come up with sound IDS, it is evident that the differentiation between normal and anomalous traffic is of utmost importance. It can therefore be concluded that in as much as it requires proper tuning and application, Logistic Regression can effectively be used as one of the components of such systems since the level of accuracy achieved in this study is very high. Moreover, the model is simple enough for implementation, thus, can be implemented in complex environments with issues concerning computational capabilities.

## 4.6 Limitation of the Present Study and Possible Directions for Further Research

But there are some limitations in the present study, which can be discussed as follows. The results which were obtained using the presented model and algorithm are outstanding; however, this model and algorithm was trained only with one set of data. Therefore, it is not yet known whether it can be used on other forms of network traffic or on other networks. The future work could be to test the viability of the model on the other datasets and traffic data collected in real time. Moreover, investigating richer paradigm is possible, for instance deep learning based paradigms, which may enable steering to other forms of patterns and increase classification quality. Lastly, the incorporation of feature engineering tools and processes for creating new, possibly more informative features also adds to the improvement of the model.

## 4.7 Conclusions

In summary the logistic regression model developed in this study has demonstrated outstanding performance in classifying network traffic data achieving a near perfect accuracy the carefully selected features combined with rigorous preprocessing and model tuning contributed to this success however while the results are promising further validation and exploration of more complex models are necessary to ensure the robustness and generalizability of the findings the insights gained from this study pave the way for future research and practical applications in the field of cybersecurity particularly in enhancing the effectiveness of intrusion detection systems.

# Chapter 5 Conclusion and Future Works

## 5. 1 Introduction

Finally in this last chapter, we synchronize the findings that have been identified from our research on the adaptive encryption scheme for IoT sensor networks. The goal of the research was to create a new algorithm for encryption that would suit the IoT best as the level of data protection can be rather diverse, while amounts of resources are often limited. To pursue the best feature of security and system performance, we tried to use the machine learning approach to classify dynamic data and use different encryption levels for them. The last chapter of the dissertation will be used to provide a conclusion that warrants the results obtained in this study, make some recommendations to future studies in enhancing IoT security.

## 5. 2 Conclusion

The research conducted in this thesis was useful to prove the feasibility of an adaptive encryption scheme for an IoT sensor networks. As a result of a machine learning algorithm to perform data classification to determine the level of data criticality used in the model, a dual layer encryption strategy where XOR encryption was used for low critical data and homomorphic encryption for the high critical data was implemented. This adaptive method enhanced the compromise between security and processing time greatly, prompting it for the utilization in IoT community that tends to experience strangled resources. Let it be recalled that the N-BaIoT dataset was used to test the scheme, and the findings confirmed its effectiveness in minimizing computational intensity even as encryption of end-user information is significantly enhanced—an aspect that goes to the heart of IoT security.

## 5. 3 Future Work

Yet, this research has offered some groundwork for adaptive encryption in the IoT network; there are a few factors that require further analysis to improve the scheme’s reliability and scalability. Further work to enhance the effectiveness of the proposed approach can be further enhancement of machine learning classification, including the integration of deep learning that can provide for higher accuracy in the classification of the data according to the established criticality levels. Furthermore, broadening the coverage of the scheme covering more IoT devices and communication protocols would also be useful in the evaluation of the capability of the scheme in terms of IoT real-world applications.

One of the possible future developments is linked to the integration of the previously described adaptive encryption scheme with other promising IT safety solutions, for example, blockchain solutions, to design more stable IoT security solutions. Analyzing this scheme for actual-time operation in the live IoT networks especially in the industrial and critical infrastructure domains, the various issues and merits could be studied. However, due to growth of emerging threats and innovations in the encryption techniques, it will be crucial to constantly monitor the adaptable scheme in order to guarantee the efficiency of the scheme as it evolves to meet the challenges facing the cybersecurity systems. With these future initiatives in mind, the idea is to extend the position of the adaptive encryption approach as one of the core security methods for the constantly growing IoT environment.

# References

Li, C. and Palanisamy, B. 2019. Privacy in Internet of Things: From Principles to Technologies. *IEEE Internet of Things Journal* 6(1), pp. 488–505. doi: <https://doi.org/10.1109/JIOT.2018.2864168>

Al-Rubaie , M. and J. Morris, C. 2019. Privacy-Preserving Machine Learning: Threats and Solutions. *IEEE Security & Privacy* 17(2), pp. 49–58. doi: <https://doi.org/10.1109/MSEC.2018.2888775>

Meurisch, C. and Mühlhäuser, M. 2021. Data Protection in AI Services. *ACM Computing Surveys* 54(2), pp. 1–38. doi: <https://doi.org/10.1145/3440754>

Ghayyur, S., Pappachan, P., Wang, G., Mehrotra, S. and Venkatasubramanian, N. 2020. Designing privacy preserving data sharing middleware for internet of things. *Proceedings of the Third Workshop on Data: Acquisition To Analysis* *(DATA '20). Association for Computing Machinery, New York, NY, USA,* *2020*. doi: <https://doi.org/10.1145/3419016.3431484>

Meneghello, F., Calore, M., Zucchetto, D., Polese, M. and Zanella, A. 2019. IoT: Internet of Threats? A survey of practical security vulnerabilities in real IoT devices. *IEEE Internet of Things Journal* 6(5), pp. 1–1. doi: <https://doi.org/10.1109/jiot.2019.2935189>

Farokhi, F. and Kâafar, M.A. 2020. Modelling and Quantifying Membership Information Leakage in Machine Learning. *arXiv preprint*. doi: <https://doi.org/10.48550/arxiv.2001.10648>

Zheng, S., Apthorpe, N., Chetty, M. and Feamster, N. 2018. User Perceptions of Smart Home IoT Privacy. *Proceedings of the ACM on Human-Computer Interaction* 2(CSCW). Association for Computing Machinery*, New York, NY, USA, 2018.* doi: <https://doi.org/10.1145/3274469>

Tawalbeh, L., Muheidat, F., Tawalbeh, M. and Quwaider, M. 2020. IoT Privacy and Security: Challenges and Solutions. *Applied Sciences* 10(12). doi: <https://doi.org/10.3390/app10124102>

Nguyen, T.M. and Vu, X.-S. 2023. Privacy and Trust in IoT Ecosystems with Big Data: A Survey of Perspectives and Challenges. In: *2023 IEEE Ninth International Conference on Big Data Computing Service and Applications (BigDataService), Athens, Greece, 2023*. doi: <https://doi.org/10.1109/bigdataservice58306.2023.00045>

Abadie, A. 2021. Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature* 59(2), pp. 391–425. doi: <https://doi.org/10.1257/jel.20191450>

Heurix, J., Zimmermann, P., Neubauer, T. and Fenz, S. 2015. A taxonomy for privacy enhancing technologies. *Computers & Security* 53, pp. 1–17. doi: <https://doi.org/10.1016/j.cose.2015.05.002>

Pfitzmann, A. and Hansen, M. 2006. *Anonymity, Unlinkability, Unobservability, Pseudonymity, and Identity Management -A Consolidated Proposal for Terminology*. Available at: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=76296a3705d32a16152875708465c136c70fe109> [Accessed: 16 July 2024].

Abdul, M. and Seong Oun, H. 2023. When AI Meets Information Privacy: The Adversarial Role of AI in Data Sharing Scenario. *IEEE Access* 11. doi: <https://doi.org/10.1109/ACCESS.2023.3297646>

Warren, S.D. and Brandeis, L.D. 1890. The Right to Privacy. *Harvard Law Review* *4*(5), p.193.

Westin, A. 1968. Washington and Lee Law Review Privacy And Freedom. Available at: <https://scholarlycommons.law.wlu.edu/cgi/viewcontent.cgi?article=3659&context=wlulr> [Accessed: 20 July 2024].

Council of Europe. 2010. *The European Convention on Human Rights ROME 4 November 1950 and its Five Protocols*. Available at: <https://new.hamoked.org.il/files/2010/5370_eng.pdf> [Accessed: 26 June 2024].

Avizienis, A., Laprie, J.-C., Randell, B. and Landwehr, C. 2004. Basic concepts and taxonomy of dependable and secure computing. *IEEE Transactions on Dependable and Secure Computing* 1(1), pp. 11–33. doi: <https://doi.org/10.1109/tdsc.2004.2>

Elahi, H., Wang, G., Peng, T. and Chen, J. 2019. AI and Its Risks in Android Smartphones: A Case of Google Smart Assistant. *Communications in computer and information science1123*. doi: <https://doi.org/10.1007/978-981-15-1304-6_27>

Danezis, G., Domingo-Ferrer, J., Hansen, M., Hoepman, J.-H., Metayer, D.L., Tirtea, R. and Schiffner, S. 2015. Privacy and Data Protection by Design - from policy to engineering. *arXiv preprint.* doi: <https://arxiv.org/abs/1501.03726>

Zagi, L.M. and Aziz, B. 2020. Privacy Attack On IoT: a Systematic Literature Review. In: *2020 International Conference on ICT for Smart Society (ICISS)*. Bandung; Indonesia, 2020. doi: <https://doi.org/10.1109/ICISS50791.2020.9307568>

Tao, H., Bhuiyan, M.Z.A., Abdalla, A.N., Hassan, M.M., Zain, J.M. and Hayajneh, T. 2019. Secured Data Collection With Hardware-Based Ciphers for IoT-Based Healthcare. *IEEE Internet of Things Journal* 6(1), pp. 410–420. doi: <https://doi.org/10.1109/jiot.2018.2854714>

Pinchot, J. and Cellante, D. 2021. Privacy Concerns and Data Sharing Habits of Personal Fitness Information Collected via Activity Trackers. *Journal of Information Systems Applied Research* 14(2). Available at: <http://jisar.org/2021-14/n2/JISARv14n2.pdf#page=4> [Accessed: 29 June 2024].

Shastri, S., Wasserman, M. and Chidambaram, V. 2019. *The Seven Sins of {Personal-Data} Processing Systems under {GDPR}*.In 11th USENIX Workshop on Hot Topics in Cloud Computing (HotCloud 19). Available at: <https://www.usenix.org/conference/hotcloud19/presentation/shastri> [Accessed: 28 June 2024].

Zhang, C., Shahriar, H. and Riad, A.B.M.K. 2020. Security and Privacy Analysis of Wearable Health Device. *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*. Madrid; Spain, 2020. doi: <https://doi.org/10.1109/COMPSAC48688.2020.00044%7D>

Dwyer, C. 2011. Privacy in the Age of Google and Facebook. *IEEE Technology and Society Magazine* 30(3), pp. 58–63. doi: <https://doi.org/10.1109/MTS.2011.942309>

de Montjoye, Y.-A., Hidalgo, C.A., Verleysen, M. and Blondel, V.D. 2013. Unique in the Crowd: The privacy bounds of human mobility. *Scientific Reports* 3(1). doi: <https://doi.org/10.1038/srep01376>

Dilmaghani, S., Brust, M.R., Danoy, G., Cassagnes, N., Pecero, J. and Bouvry, P. 2019. Privacy and Security of Big Data in AI Systems: A Research and Standards Perspective. In: *2019 IEEE International Conference on Big Data (Big Data)*. Los Angeles, CA, USA, 2019. doi: <https://doi.org/10.1109/BigData47090.2019.9006283>

Sweeney, L. 2000. *Simple Demographics Often Identify People Uniquely*. Carnegie Mellon University, Data Privacy Working Paper 3. Available at: <https://privacytools.seas.harvard.edu/sites/projects.iq.harvard.edu/files/privacytools/files/paper1.pdf>.

Shokri, R. and Shmatikov, V. 2015. Privacy-Preserving Deep Learning. In: *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security - CCS ’15*. New York, NY, USA, 2015. doi: <https://doi.org/10.1145/2810103.2813687>

Fredrikson, M., Jha, S. and Ristenpart, T. 2015. Model Inversion Attacks that Exploit Confidence Information and Basic Countermeasures. *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security - CCS ’15*. New York, NY, USA, 2015. doi: https://doi.org/10.1145/2810103.2813677

Wang, Y., Pan, Y., Yan, M., Su, Z. and Luan, T. 2023. A Survey on ChatGPT: AI-Generated Contents, Challenges, and Solutions. *IEEE Open Journal of the Computer Society* 4. doi: 10.1109/OJCS.2023.3300321

Edu, J.S., Such, J.M. and Suarez-Tangil, G. 2020. Smart Home Personal Assistants. *ACM Computing Surveys* 53(6), pp. 1–36. doi: <https://doi.org/10.1145/3412383>

Lau, J., Zimmerman, B. and Schaub, F. 2018. Alexa, Are You Listening? Privacy Perceptions, Concerns and Privacy-seeking Behaviors with Smart Speakers. *Proceedings of the ACM on Human-Computer Interaction* 2(CSCW), pp. 1–31. doi: <https://doi.org/10.1145/3274371>

Smith, M. and Miller, S. 2021. The ethical application of biometric facial recognition technology. *AI & SOCIETY* 37(1), pp. 167–175. doi: <https://doi.org/10.1007/s00146-021-01199-9>

Chaum, D.L. 1981. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*24(2), pp. 84–90. doi: <https://doi.org/10.1145/358549.358563>

Cha, S.-C., Hsu, T.-Y., Xiang, Y. and Yeh, K.-H. 2018. Privacy Enhancing Technologies in the Internet of Things: Perspectives and Challenges. *IEEE Internet of Things Journal*, pp. 1–1. doi: <https://doi.org/10.1109/jiot.2018.2878658>

Ziegeldorf, J.H., Morchon, O.G. and Wehrle, K. 2013. Privacy in the Internet of Things: threats and challenges. *Security and Communication Networks* 7(12), pp. 2728–2742. doi: <https://doi.org/10.1002/sec.795>

Ustundag Soykan, E., Karacay, L., Karakoc, F. and Tomur, E. 2022. A Survey and Guideline on Privacy Enhancing Technologies for Collaborative Machine Learning. *IEEE Access* 10, pp. 97495–97519. doi: <https://doi.org/10.1109/access.2022.3204037>

Dwork, C. 2006. Differential Privacy. *Automata, Languages and Programming* 4052, pp. 1–12. doi: <https://doi.org/10.1007/11787006_1>

Dwork, C. and Roth, A. 2014. The Algorithmic Foundations of Differential Privacy. *Foundations and Trends® in Theoretical Computer Science* 9(3-4), pp. 211–407. doi: <https://doi.org/10.1561/0400000042>

Zhu, T., Ye, D., Wang, W., Zhou, W. and Yu, P. 2021. More Than Privacy: Applying Differential Privacy in Key Areas of Artificial Intelligence. *IEEE Transactions on Knowledge and Data Engineering* 34(6), pp. 1–1. doi: <https://doi.org/10.1109/tkde.2020.3014246>

Xu, C., Ren, J., She, L., Zhang, Y., Qin, Z. and Ren, K. 2019a. EdgeSanitizer: Locally Differentially Private Deep Inference at the Edge for Mobile Data Analytics. *IEEE Internet of Things Journal* 6(3), pp. 5140–5151. doi: <https://doi.org/10.1109/jiot.2019.2897005>

|  |
| --- |
| Abadi, M., Chu, A., Goodfellow, I., McMahan, H.B., Mironov, I., Talwar, K. and Zhang, L. 2016. Deep Learning with Differential Privacy. In: *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security - CCS’16 (ACM CCS*), pp. 308-318. doi: [https://doi.org/10.48550/arXiv.1607.00133](%20https://doi.org/10.48550/arXiv.1607.00133) |

Yang, M., Zhu, T., Liu, B., Xiang, Y. and Zhou, W. 2018. Machine Learning Differential Privacy With Multifunctional Aggregation in a Fog Computing Architecture. *IEEE Access* 6, pp. 17119–17129. doi: <https://doi.org/10.1109/ACCESS.2018.2817523>

Rivest, R.L., Shamir, A. and Adleman, L. 1977. *A Method for Obtaining Digital Signatures and Public-Key Cryptosystems*. Available at: <https://people.csail.mit.edu/rivest/Rsapaper.pdf>. [Accessed: 28 July 2024].

Paillier, P. 1999. *Public-Key Cryptosystems Based on Composite Degree Residuosity Classes.*. Available at: https://link.springer.com/content/pdf/10.1007/3-540-48910-X\_16.pdf [Accessed: 25 July 2024].

Gentry, C. 2009. Fully homomorphic encryption using ideal lattices. *Proceedings of the 41st annual ACM symposium on Symposium on theory of computing - STOC ’09*. doi: <https://doi.org/10.1145/1536414.1536440>

Li, S., Zhao, S., Min, G., Qi, L. and Liu, G. 2021. Lightweight Privacy-Preserving Scheme using Homomorphic Encryption in Industrial Internet of Things. *IEEE Internet of Things Journal*. doi: <https://doi.org/10.1109/jiot.2021.3066427>

Huang, Y. and Duan, G. 2024. A Privacy-Preserving Decision Tree Evaluation Scheme for Multiple Wearable Devices. In: *2024 27th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*. Tianjin, China. doi: <https://doi.org/10.1109/cscwd61410.2024.10580704>

Bautista, O.G. and Akkaya, K. 2022. Network-Efficient Pipelining-Based Secure Multiparty Computation for Machine Learning Applications. In: *2022 IEEE 47th Conference on Local Computer Networks (LCN)*. Edmonton, AB, Canada. doi: <https://doi.org/10.1109/lcn53696.2022.9843372>

Tonyali, S., Akkaya, K., Saputro, N., Uluagac, A.S. and Nojoumian, M. 2018. Privacy-preserving protocols for secure and reliable data aggregation in IoT-enabled Smart Metering systems. *Future Generation Computer Systems* 78, pp. 547–557. doi: <https://doi.org/10.1016/j.future.2017.04.031>

Jarin, I. and Eshete, B. 2021. PRICURE: Privacy-Preserving Collaborative Inference in a Multi-Party Setting. *arXiv (Cornell University)*. doi: <https://doi.org/10.1145/3445970.3451156>

McMahan H, B., Moore, E., Ramage, D., Hampson, S. and Arcas, B.A. y. 2016. Communication-Efficient Learning of Deep Networks from Decentralized Data. In: *Artificial intelligence and statistics*. Available at: <https://arxiv.org/abs/1602.05629> [Accessed: 4 July 2024].

Konečný, J., McMahan, H.B., Ramage, D. and Richtárik, P. 2016. Federated Optimization: Distributed Machine Learning for On-Device Intelligence. *arXiv preprint*. doi: <https://arxiv.org/abs/1610.02527>.

Konečný, J., McMahan, H.B., Yu, F.X., Richtárik, P., Suresh, A.T. and Bacon, D. 2017. Federated Learning: Strategies for Improving Communication Efficiency. *arXiv preprint*. doi: <https://arxiv.org/abs/1610.05492>

Truex, S., Baracaldo, N., Anwar, A., Steinke, T., Ludwig, H., Zhang, R. and Zhou, Y. 2019. A Hybrid Approach to Privacy-Preserving Federated Learning. *Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security - AISec’19*. doi: <https://doi.org/10.1145/3338501.3357370>

Xu, R., Baracaldo, N., Zhou, Y., Anwar, A. and Ludwig, H. 2019b. HybridAlpha. In: *Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security - AISec’19*. doi: <https://doi.org/10.1145/3338501.3357371>

Singh, S., Rathore, S., Alfarraj, O., Tolba, A. and Yoon, B. 2021. A framework for privacy-preservation of IoT healthcare data using Federated Learning and blockchain technology. *Future Generation Computer Systems*. doi: <https://doi.org/10.1016/j.future.2021.11.028>

Alaya, B., Laouamer, L. and Msilini, N. 2020. Homomorphic encryption systems statement: Trends and challenges. *Computer Science Review* 36, p. 100235. doi: <https://doi.org/10.1016/j.cosrev.2020.100235>

Feng, Y., Bai, T., Lu, S., Tang, X. and Wu, J. 2023. SMPC Task Decomposition: A Theory for Accelerating Secure Multi-party Computation Task. *arXiv.org*. doi: <https://doi.org/10.48550/arXiv.2303.00343>

Hard, A. et al. 2019. Federated Learning for Mobile Keyboard Prediction. *arXiv preprint*. doi: <https://arxiv.org/abs/1811.03604>

Princi, E. and Krämer, N.C. 2020. Out of control – privacy calculus and the effect of perceived control and moral considerations on the usage of IoT healthcare devices. *Frontiers in Psychology* 11. Available at: <https://doi.org/10.3389/fpsyg.2020.582054>.

Morel, V. and Fischer-Hübner, S. 2023. Automating privacy decisions -where to draw the line? *cs.CY*. doi: <https://arxiv.org/pdf/2305.08747>.

Du, X., Chen, H.-H., Zhu, L., Li, J. and Chang, Z. 2018. Security and Privacy in Wireless IoT. *IEEE Wireless Communications* 25(6), pp. 10–11. doi: <https://doi.org/10.1109/mwc.2018.8600750>.

Andreotta, A.J., Kirkham, N. and Rizzi, M. 2021. AI, big data, and the future of consent. *AI & SOCIETY* 37. doi: <https://doi.org/doi.org/10.1007/s00146-021-01262-5>.

Gerke, S., Minssen, T. and Cohen, I.G. 2020. *Ethical and Legal Challenges of Artificial Intelligence-Driven Healthcare*. Available at: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3570129>.

Bae, H., Jang, J., Jung, D., Jang, H., Ha, H. and Yoon, S. 2018. Security and Privacy Issues in Deep Learning. *ArXiv*. doi: arxiv:1807.11655v2.

GDPR.EU. 2023. *General Data Protection Regulation (GDPR) Compliance Guidelines*. Available at: [https://gdpr.eu](https://gdpr.eu/) [Accessed: 28 June 2024].

Menard, S. 2002. *Applied Logistic Regression Analysis*. SAGE.

Meidan, Y., Bohadana, M., Mathov, Y., Mirsky, Y., Shabtai, A., Breitenbacher, D. and Elovici, Y. 2018. N-BaIoT—Network-Based Detection of IoT Botnet Attacks Using Deep Autoencoders. *IEEE Pervasive Computing* 17(3), pp. 12–22. doi: https://doi.org/10.1109/mprv.2018.03367731.

# Appendix:

# %% [code] {"execution":{"iopub.status.busy":"2024-08-15T16:02:35.367344Z","iopub.execute\_input":"2024-08-15T16:02:35.367854Z","iopub.status.idle":"2024-08-15T16:02:50.932441Z","shell.execute\_reply.started":"2024-08-15T16:02:35.367814Z","shell.execute\_reply":"2024-08-15T16:02:50.931005Z"}}

!pip install tenseal

# %% [markdown]

#

# This code processes a dataset containing benign and malicious network traffic data to classify its criticality.

# It trains a Logistic Regression model to predict whether data points are low-critical (benign) or high-critical (malicious) based on selected features.

# Then, an adaptive encryption function encrypts the data points in the test set: low-critical data is encrypted using XOR encryption, while high-critical data is encrypted using a homomorphic encryption-like method.

# The code use batch processing by encrypting the test data in chunks becuase we are not using GPU

# %% [markdown]

# # Loadind Dataset

# %% [code]

# %% [code] {"execution":{"iopub.status.busy":"2024-08-15T16:02:50.935278Z","iopub.execute\_input":"2024-08-15T16:02:50.935714Z","iopub.status.idle":"2024-08-15T16:02:55.508188Z","shell.execute\_reply.started":"2024-08-15T16:02:50.935664Z","shell.execute\_reply":"2024-08-15T16:02:55.506783Z"}}

import pandas as pd

from sklearn.model\_selection import train\_test\_split

from sklearn.preprocessing import StandardScaler

from sklearn.linear\_model import LogisticRegression

import numpy as np

# Load and preprocess the dataset

benign = pd.read\_csv('/kaggle/input/nbaiot-dataset/1.benign.csv')

malicious = pd.read\_csv('/kaggle/input/nbaiot-dataset/1.mirai.syn.csv')

# Combine the datasets and assign labels

benign['label'] = 0

malicious['label'] = 1

data = pd.concat([benign, malicious], ignore\_index=True)

# Feature selection

features = ['MI\_dir\_L5\_weight', 'MI\_dir\_L5\_mean', 'MI\_dir\_L5\_variance',

'MI\_dir\_L3\_weight', 'MI\_dir\_L3\_mean', 'MI\_dir\_L3\_variance']

X = data[features]

y = data['label']

# %% [markdown]

# \*\*Data visualization before training and testing\*\*

# %% [code] {"execution":{"iopub.status.busy":"2024-08-15T16:02:55.509801Z","iopub.execute\_input":"2024-08-15T16:02:55.510218Z","iopub.status.idle":"2024-08-15T16:03:03.414526Z","shell.execute\_reply.started":"2024-08-15T16:02:55.510186Z","shell.execute\_reply":"2024-08-15T16:03:03.413203Z"}}

# Combine X and y for easier plotting

plot\_data = X.copy()

plot\_data['label'] = y

# Set up the matplotlib figure

plt.figure(figsize=(15, 10))

# Plot each feature distribution by label

for i, feature in enumerate(features, 1):

plt.subplot(2, 3, i)

sns.histplot(data=plot\_data, x=feature, hue='label', kde=True, element='step', stat='density')

plt.title(f'Distribution of {feature} by Class')

plt.xlabel(feature)

plt.ylabel('Density')

plt.tight\_layout()

plt.show()

# %% [code] {"execution":{"iopub.status.busy":"2024-08-15T16:03:03.416062Z","iopub.execute\_input":"2024-08-15T16:03:03.416451Z","iopub.status.idle":"2024-08-15T16:03:03.955258Z","shell.execute\_reply.started":"2024-08-15T16:03:03.416417Z","shell.execute\_reply":"2024-08-15T16:03:03.953565Z"}}

# Split the dataset

x\_train, x\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

# Standardize the features

scaler = StandardScaler()

x\_train\_scaled = scaler.fit\_transform(x\_train)

x\_test\_scaled = scaler.transform(x\_test)

# Train the Logistic Regression model

model = LogisticRegression(random\_state=0)

model.fit(x\_train\_scaled, y\_train)

# Calculate accuracy on the test set

accuracy = model.score(x\_test\_scaled, y\_test)

print(f"Accuracy of the Logistic Regression model: {accuracy:.4f}")

# %% [markdown]

# # Adaptive encryption

# %% [code] {"execution":{"iopub.status.busy":"2024-08-15T16:03:03.960353Z","iopub.execute\_input":"2024-08-15T16:03:03.961429Z","iopub.status.idle":"2024-08-15T16:03:26.211492Z","shell.execute\_reply.started":"2024-08-15T16:03:03.961361Z","shell.execute\_reply":"2024-08-15T16:03:26.210070Z"}}

# Efficient encryption functions

def xor\_encrypt(data):

return ''.join(chr(ord(char) ^ 42) for char in str(data)) #uses a hardcoded key in the XOR encryption function which is 42.

def homomorphic\_encrypt(data):

# Simulating encryption by converting data to a string with a prefix

return f"homomorphic\_encrypted\_{hash(str(data))}"

# Adaptive encryption function with batch processing

def adaptive\_encryption(data, model, scaler, features, batch\_size=50):

encrypted\_data = []

for start in range(0, len(data), batch\_size):

end = min(start + batch\_size, len(data))

batch = data.iloc[start:end]

# Predict the criticality of the batch

scaled\_features = scaler.transform(batch[features])

predictions = model.predict(scaled\_features)

for i, prediction in enumerate(predictions):

if prediction == 0: # Low-critical

encrypted\_data.append(xor\_encrypt(batch.iloc[i].to\_string()))

else: # High-critical

encrypted\_data.append(homomorphic\_encrypt(batch.iloc[i].to\_string()))

return encrypted\_data

# Run adaptive encryption on the test set

encrypted\_data = adaptive\_encryption(x\_test, model, scaler, features, batch\_size=50)

# Display some encrypted results

for i, encrypted in enumerate(encrypted\_data[:5]):

print(f"Encrypted data point {i+1}: {encrypted}")

# %% [markdown]

# the encryption simulation for high-critical data is working as expected, producing unique "homomorphic encrypted" strings based on the data's hash values. These strings represent the encrypted form of

# data points, simulating the behavior of a real homomorphic encryption scheme.

# %% [markdown]

# # Validation of the Process

# %% [code] {"execution":{"iopub.status.busy":"2024-08-15T16:03:26.213093Z","iopub.execute\_input":"2024-08-15T16:03:26.213466Z","iopub.status.idle":"2024-08-15T16:03:26.245092Z","shell.execute\_reply.started":"2024-08-15T16:03:26.213438Z","shell.execute\_reply":"2024-08-15T16:03:26.243642Z"}}

def validate\_encryption\_process(data, labels, model, scaler, features):

correct\_classification = 0

incorrect\_classification = 0

# Convert DataFrame to NumPy array

data\_array = data[features].values

# Iterate over data and labels

for i in range(len(data)):

scaled\_features = scaler.transform([data\_array[i]])

prediction = model.predict(scaled\_features)[0]

# Compare with true label

if prediction == labels.iloc[i]:

correct\_classification += 1

else:

incorrect\_classification += 1

# Check encryption type applied

if prediction == 0: # low-critical, XOR encryption

encrypted = xor\_encrypt(data.iloc[i].to\_string())

else: # high-critical, homomorphic encryption

encrypted = homomorphic\_encrypt(data.iloc[i].to\_string())

print(f"Data point {i}: Classified as {prediction}, Encrypted as {encrypted[:20]}...") # Print first 20 chars

return correct\_classification, incorrect\_classification

# Run validation on a small subset for simplicity

data\_subset = x\_test.head(10).reset\_index(drop=True)

labels\_subset = y\_test.head(10).reset\_index(drop=True)

correct, incorrect = validate\_encryption\_process(data\_subset, labels\_subset, model, scaler, features)

print(f"\nCorrect Classifications: {correct}")

print(f"Incorrect Classifications: {incorrect}")

# %% [markdown]

# # Data visulaization

# %% [code] {"execution":{"iopub.status.busy":"2024-08-15T16:03:26.246636Z","iopub.execute\_input":"2024-08-15T16:03:26.247132Z","iopub.status.idle":"2024-08-15T16:03:26.268969Z","shell.execute\_reply.started":"2024-08-15T16:03:26.247086Z","shell.execute\_reply":"2024-08-15T16:03:26.267271Z"}}

y\_test\_pred = model.predict(x\_test\_scaled)

# %% [code] {"execution":{"iopub.status.busy":"2024-08-15T16:03:26.271358Z","iopub.execute\_input":"2024-08-15T16:03:26.272020Z","iopub.status.idle":"2024-08-15T16:03:26.629651Z","shell.execute\_reply.started":"2024-08-15T16:03:26.271963Z","shell.execute\_reply":"2024-08-15T16:03:26.628212Z"}}

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn.metrics import confusion\_matrix, ConfusionMatrixDisplay

# Generate predictions

y\_test\_pred = model.predict(x\_test\_scaled)

# Calculate the confusion matrix

conf\_matrix = confusion\_matrix(y\_test, y\_test\_pred)

# Plot using seaborn heatmap

plt.figure(figsize=(8, 6))

sns.heatmap(conf\_matrix, annot=True, fmt='d', cmap='Blues',

xticklabels=['Low-Critical', 'High-Critical'],

yticklabels=['Low-Critical', 'High-Critical'])

plt.title('Confusion Matrix')

plt.xlabel('Predicted')

plt.ylabel('Actual')

plt.show()

# %% [markdown]

# \* True Positives (TP): 24,563 (High-Critical predicted correctly)

# \* True Negatives (TN): 9,852 (Low-Critical predicted correctly)

# \* False Positives (FP): 0 (No Low-Critical data points were incorrectly predicted as High-Critical)

# \* False Negatives (FN): 10 (A small number of High-Critical data points were incorrectly predicted as Low-Critical)

# The confusion matrix indicates that the model is highly accurate, with very few misclassifications, particularly in distinguishing between Low-Critical and High-Critical data points.

# %% [code] {"execution":{"iopub.status.busy":"2024-08-15T16:03:26.631430Z","iopub.execute\_input":"2024-08-15T16:03:26.631903Z","iopub.status.idle":"2024-08-15T16:03:27.000812Z","shell.execute\_reply.started":"2024-08-15T16:03:26.631867Z","shell.execute\_reply":"2024-08-15T16:03:26.999316Z"}}

# ROC Curve and AUC

from sklearn.metrics import roc\_curve, auc

y\_test\_prob = model.predict\_proba(x\_test\_scaled)[:, 1]

fpr, tpr, \_ = roc\_curve(y\_test, y\_test\_prob)

roc\_auc = auc(fpr, tpr)

plt.figure(figsize=(8, 6))

plt.plot(fpr, tpr, color='blue', label=f'ROC curve (area = {roc\_auc:.2f})')

plt.plot([0, 1], [0, 1], color='gray', linestyle='--')

plt.xlim([0.0, 1.0])

plt.ylim([0.0, 1.05])

plt.xlabel('False Positive Rate')

plt.ylabel('True Positive Rate')

plt.title('Receiver Operating Characteristic (ROC) Curve')

plt.legend(loc='lower right')

plt.show()

# %% [markdown]

# The ROC curve is a plot of the True Positive Rate (TPR) versus the False Positive Rate (FPR).

# the curve closely hugs the top-left corner of the graph, which is an ideal scenario, indicating a strong performance by the model.

# '

# Area Under the Curve (AUC): The value is 1.00, which is the best possible score, showing that the model is perfectly distinguishing between the classes.

# %% [code] {"execution":{"iopub.status.busy":"2024-08-15T16:03:27.002593Z","iopub.execute\_input":"2024-08-15T16:03:27.003132Z","iopub.status.idle":"2024-08-15T16:03:28.025039Z","shell.execute\_reply.started":"2024-08-15T16:03:27.003077Z","shell.execute\_reply":"2024-08-15T16:03:28.022854Z"}}

# Feature Importance (for Logistic Regression, if applicable)

if hasattr(model, 'coef\_'):

feature\_importance = pd.Series(model.coef\_[0], index=features).sort\_values(ascending=False)

plt.figure(figsize=(8, 6))

feature\_importance.plot(kind='bar')

plt.title('Feature Importance')

plt.ylabel('Coefficient Value')

plt.show()