National College of Ireland

Enhanced Security for Insecure Systems within Zero Trust Architecture

MSc Research Project
Cybersecurity

Pradeep Prakash
Student ID: 21215413

School of Computing
National College of Ireland

Supervisor: Evgeniia Jayasekera

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: Pradeep Prakash

Student ID: 21215413

Programme: MSc in Cybersecurity **Year:** 2022-2023

Module: MSc Research Project

Lecturer: Evgeniia Jayasekera

Submission Due Date: 14-Aug-2023

Project Title: Enhanced Security for Insecure Systems within Zero trust Architecture

Word Count: 6670 **Page Count:** 20

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Pradeep Prakash

Date: 14-08-2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Enhanced Security for Insecure Systems within ZTA

Pradeep Prakash
21215413

Abstract

The research work investigated different concepts related to the "Zero-Trust Architecture (ZTA)" and its application in increasing the security of APIs. The first section of this paper identifies the key aim of this research work with the development of key objectives of this investigation. The second section of this paper developed a critical analysis of previous research work with the identification of their limitations and advantages. Thus, the paper develops some critical research gaps and develops research questions as per the research gaps. The third part of this research work investigates different methods considered in the research with the identification of different frameworks. Users' identification is developed with the application of "least access policies" with multiple authentications for improving the authentication process with increasing network security. Thus, this analysis comprises different considered tools and techniques with ethical considerations in the investigation. The seventh part of this research work consist of conclusion and discussion including recommendations and key findings. Classifiers can rely on each recommending area of assessment using vast communicative process management in order to be more adequate in nature. Addressing in each estimation value overall securable approaches are designed in the software areas of Jupyter Notebook or Google colab using Python computation language.

Keywords: "Zero-Trust Architecture (ZTA)", "Cyber Attack Exploitation", "Trust Management System", "Blockchain", "Authorization", "Internet of Things (IoT)", "application programming interface (API)".

1. Introduction:

The development of networks is increasing with the increasing rate of cyber-attacks creating threats to the data system of a network. The application of traditional perimeter increases the risk of cyber security that requires to be developed with the application of better visibility for controlling access to the data system. The main concern associated with data management in different organizations is data security due to the increasing number of cyber-attacks in different organizations. This network accessibility within an organization can be controlled by developing a "Zero-trust security" architecture for developing strict accessibility capabilities to different people and connected devices with the network. Moreover, only 11% of global organizations have incorporated security policy with their "application programming interfaces (APIs)" which increases protection risk in other organizations. Thus, the development of network security by controlling access to the data system is necessary. Moreover, the identification of the users and ensuring their security concerns with the user devices are necessary to develop for securing the network with authentication.

The application of "Zero-Trust Architecture (ZTA)" secures different types of vulnerable systems connected to the network such as "phishing, stolen credentials, and out-of-date devices. Zero trust is implemented and developed for making the authentication process strong for both the users and the application to minimize the incorporation of policy access. Zero trust helps in monitoring and gaining access to controls in order to keep the data, resources, and information safe from the attackers. It helps in identifying different models of already applied ZTA and identifies the design principles. This course work is much more significant as the aim of the vulnerability of data from exploiting insecure APIs under secured conditions was much more emphasized. The already existing API security issues may be assessed for the designation of their effect on ZTA applications.

Existing technological models may be investigated to get all possible modifications in the ZTA application along with the growth of a "zero-trust environment (ZTA)" in order to APIs (Leahy and Thorpe, 2022). The comprehensive application of "Zero-Trust Architecture (ZTA)" is developed through three main factors such as "Users, Applications, and Infrastructure". Moreover, the application of "Zero Trust" removes all types of implicit trust in the application with the concept that the application is not trusted and requires continuous authentication. Everything connected to the infrastructure of the network such as "routers, switches, cloud, IoT, and supply chain" is addressed through "Zero trust".

1.1 Research aim

This research work aims to develop an application approach and process of "Zero-Trust Architecture (ZTA)" for securing vulnerable "application programming interfaces (APIs)" with increasing security concerns through visibility and accessibility.

1.2 Research objectives

The key objectives of this research work are:

- To examine the importance of the application of "Zero-Trust Architecture (ZTA)" in increasing the security of insecure API network.
- To determine the flexible method of application of the "Zero-Trust Architecture (ZTA)" in a large scale for increasing network security.
- To investigate the impact of the "continual verification and authorization" on the system operation and its users as well.
- To determine the compatibility level of different techniques used for security development with "Zero-Trust Architecture (ZTA)" with its appropriate framework identification.

To develop a comprehensive "migration architecture" for integration of the "Zero-Trust Architecture (ZTA)" framework with identification of its limitations.

2. Literature review

He et al. (2022) explained the different challenges of applying different technologies using the "Zero trust model" for security development along with the identification of their future trends of applications. The paper discussed the detailed architecture of the "Zero trust authentication (ZTA)" model by combining it with "artificial intelligence" for maximizing security. Thus, the paper has described different evaluation models of ZTA application and developed their comparative analysis to find different advantages and disadvantages.

Horne and Nair (2021) examined the design principles of the application of the ZTA model for developing "zero trust hype" in providing network security. The paper has represented "Zero Trust by Design (ZTBD)" with the analysis of its application principles beyond its architectural analysis. Thus, the paper has provided the trust evaluation process flow with the network architecture for improving network security.

Chen et al. (2020) analyzed the application of ZTA in protecting and creating security awareness within the 5G-based "smart healthcare system". Some key features of the 5G network have been identified in this paper such as "high bandwidth, low latency, and high concurrency". The paper has identified security requirements associated with the "5G-based smart healthcare system" for protecting its huge data satirized in the management system. key dimension of incorporating ZTA has been identified in this analysis and these are "users, terminals, and applications" for increasing real-time security.

Ameer et al. (2022) examined uncertainties associated with the application of the “Zero Trust (ZT) paradigm” for the development of the “Zero trust architecture (ZTA)”. Thus, the paper has identified vulnerabilities associated with the smart applications of IoT systems in different networks and the necessity of building zero trust within the network. The different case uses have been explored by the researchers to develop the most accurate “ZT authorization requirements framework (ZT-ARF)”.

Csikor et al. (2022) evaluated practical deployment strategies of ZTA within DNS infrastructure for developing user authentication processes. Thus, the paper has evaluated the smart communication approach with network security in DNS infrastructure.

Teerakanok et al. (2021) have examined the “challenges, steps, and things to consider” of applying legacy architecture with the application of ZTA. Thus, the paper has identified logical components required for the incorporation of the ZTA framework within network security.

Syed et al. (2022) Examined different fundamental tenets for the application of “zero trust” for the successful realization of its accuracy. The paper has identified different authentication mechanisms and their influences on the development of “zero trust” within a network.

Phiayura and Teerakanok (2023) examined the migration framework of ZTA within different organizations through the smooth application of the “Zero Trust journey”. Thus, the paper has identified challenges in migrating ZTA security within a network.

Liu et al. (2022) developed the concept of “Zero Trust Internet-of-Things” with the development of a decentralization approach to data through blockchain analysis.

Mehraj and Banday (2020) evaluated the application of a “zero trust framework” within cloud computing to improve data confidentiality with a multifaceted “cloud computing system”.

The subject of securing overall system inside zero trust architecture with ML model appears as an underexplored area in the context of the investigated literature and suggests additional research. He et al. (2022) has demonstrated the distinct challenges faced in applying additional technologies that are used for the “Zero trust model” in case of security development. It also assists in specifying additional models involving ZTA but it failed in furnishing suitable architectural conditions for the application of ZTA. Horne and Nair (2021) have inspected the strategy regulations of the application of the ZTA model for the evolution of “zero trust hype” in case of providing network security but it was unfit in providing an application procedure of ZTA. Chen et al. (2020) analyzed the application of ZTA in protecting and creating security awareness within the 5G-based “smart healthcare system” but were not able to describe the passive security standards with the ZTA application. Ameer et al. (2022) examined uncertainties that are associated with the application of the “Zero Trust (ZT) paradigm” for the development of the “Zero trust architecture (ZTA)” but were not able to specify integration techniques and technologies of ZTA that are with IoT systems.

Csikor et al. (2022) have estimated the smart transmission approach with network security in DNS infrastructure but cannot be used in other IoT-based systems. Teerakanok et al. (2021) have done an investigation of “challenges, steps, and things to consider” during the application of ZTA architecture but it cannot be used in larger networks with the migration of ZTA. Syed et al. (2022) have done a broad survey that helps in specifying crucial infrastructure of the application of ZTA. Phiayura and Teerakanok (2023) have examined the migration framework of ZTA but failed in designing an adequate framework for the ZTA application. Mehraj and Banday (2020) have investigated the integration of ZTA in cloud computing, but they failed in representing the integration of ZTA in different types of developed technologies.

The below table summarizes limitations and other metrics of the different research papers which is used for literature review:

Selected articles	**Pros**	**Limitations**	**Other metrics**
He *et al.* (2022)	Identifies different models of applying ZTA with their pros and limitations	Failed to provide the most appropriate architectural requirement for the application of ZTA	"Trust Assessment Algorithm"
Horne and Nair (2021)	Identifies design principles of ZTA	Cannot provide application approach of ZTA	"Zero Trust by Design (ZTBD)"
Chen *et al.* (2020)	Analysis Application of ZTA in smart healthcare	Cannot describe passive security measures with the ZTA application	"5-G based smart healthcare system"
Ameer *et al.* (2022)	Develops cyber security concerns with the application of ZTA	Cannot identify integration strategies and technologies of ZTA with IoT systems	"IoT-based smart system"
Csikor *et al.* (2022)	Identifies practical deployment strategies of ZTA within DNS infrastructure for developing user authentication processes.	Focuses on the DNS infrastructure cannot be applied in another IoT-based system	"ZeroDNS"
Teerakanok *et al.* (2021)	Analysis of "challenges, steps, and things to consider" during the application of ZTA architecture	This cannot be applied in larger industrial use as it is only applicable to smaller networks with the migration of ZTA	Cloud technologies

Syed et al. (2022)	Identifies critical infrastructure of the application of ZTA through a comprehensive survey	Cannot identify limitations and required components for applying ZTA	"Identity, credential, and access management (ICAM)"
Phiayura and Teerakanok (2023)	Identifies migration challenges of the ZTA framework	The application pilot program cannot develop an effective framework for the ZTA application	"ZTA migration framework"
Liu et al. (2022)	Identifies the application approach of zero trust on IoT	Cannot identify the potential impact of "trusted third-party (TTP)" on network security with IoT	"Zero Trust Internet-of-Things"
Mehraj and Banday (2020)	Analysis Integration of ZTA in cloud computing	Cannot describe the integration of ZTA in other types of advanced technologies	"Zero Trust Strategy"

Table 1: Identified limitations of previous research works.

2.1 Research Gap and research questions

Research gap 1: Depth understanding of the influence of low security of API on the architecture of "Zero trust".

Research question 1: What is the impact of the insecure APIs over the security measure of "Zero Trust architecture"?

Existing API security issues can be considered for the identification of proper answers to this research question with the identification of their impact on ZTA applications.

Researchers gap 2: Probable development for advanced applications of technologies in "Zero Trust architecture" for securing APIs.

Research question 2: What are the possible developments required for advanced applications of technologies in "Zero Trust architecture" for securing APIs?

Existing technology can be analyzed for getting all possible improvements in the application of ZTA with the development of a "zero-trust environment (ZTA)" for securing APIs.

Researchers gap 3: Comparative analysis between the security development of ZTA network with other conventional security development processes.

Research question 3: How do the security measures of ZTA network differ from other conventional applications of security measures?

This research gap can be filled with the analysis of the advantages and limitations of different security frameworks with the comparative analysis of the proposed ZTA network application in providing security to both APIs and overall system.

3. Research Methodology

3.1 Conceptual Framework

As per the overall analysis through identified secondary resources in the literature review key concepts of the research work have been considered in developing its "conceptual framework". The overall concept is developed by following the identified research questions to formulate their solution through the research. Some key concepts identified as the conceptual framework consideration in this investigation are:

- Evaluation of insecurity associated with the APIs as well as their influences on the application of ZTA.
- Evaluation of different technologies and emerging digitization concepts to improve the security of APIs with the application of ZTA.
- Comparative analysis of the performance of ZTA with other traditional approaches for developing the security of APIs.

3.2 Research Method

Analysis of the overall method for the development of the research work is necessary for the identification of required resources in the investigation. Researchers have applied "secondary research" with the considerations of "qualitative data" for the development of the research work. Researchers have applied "interpretivism philosophy" to contrast the overall framework of the research work with the identification of the other process. According to Newman and Gough (2020), this philosophical concept of researchers developed a naturalistic approach to data collection with subjective assumptions. Moreover, researchers have applied an "inductive approach" for data collection as they can identify the pattern in the dataset for further development of the analysis of their investigations. This paper is developed by following the concept of this "inductive approach" with greater flexibility to identify patterns of the available data and to form the relevant theory of application of "zero-trust architecture (ZTA)".

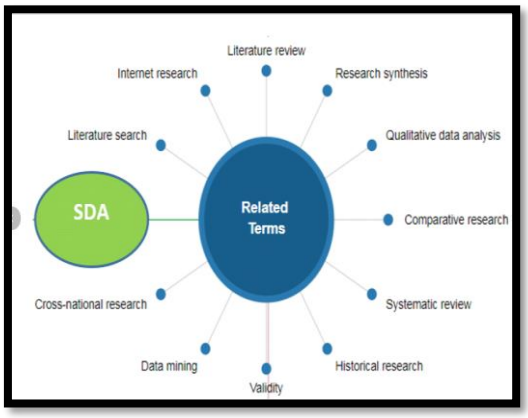

Figure 3.2: Overall concept of secondary research method

(Source: Ismoilov, 2020)

The paper has been developed by following a "descriptive design" for the development of the description of the key findings by following the data trends. This research work is developed by following the concept of "secondary research" through the development of primary research to generate the generalized concept of the application of the ZTA framework in developing network security with cloud computing (Alharahsheh and Pius, 2020). Researchers have developed statistical analysis in the literature review for the development of the current analysis. This research work is developed by following the aim of the research with the identification of the key problem associated with the security issue in APIs and its security development with the application of ZTA.

3.3 Tools and techniques used:

Consideration of different tools along with required techniques is necessary to understand the specifications of the research work and its expected outcomes. Researchers have used **secondary data** that is collected from Kaggle to generate different visualizations on the **Jupyter Notebook** or **Google Colab** with the application of **Python language**. Thus, this investigation is important for developing visualization of the developed security with the application of ZTA (Rashid *et al.* 2019). Moreover, this research work is developed by following different analysis perch through violation in Jupyter Notebook. Different techniques employed for data analysis are Data loading, Data Cleaning, Pre-Processing, Data visualization, Feature Extraction, and Data Splitting. All these techniques are applied for the visualization of the patterns of the dataset as per its different attributes and considerations in the investigation. Researchers have used Google Scholar, and ProQuest for the collection of peer-reviewed papers for investigating the application of ZTA in different security concerns with the identification of new trends of its application.

3.4 Data collection

This research work is developed with the collection of "secondary data" from different available resources. This improves the quality of the research work by encouraging researchers to further investigate the application of ZTA in the security concern of APIs. The overall investigation included a collection of different literature concepts from available research work in different "journals, and articles" by considering some key terms associated with the application of ZTA in Increasing the network security of APIs (McGill *et al.* 2021). This influences the collection of the most relevant data with qualitative value for the development of the research framework.

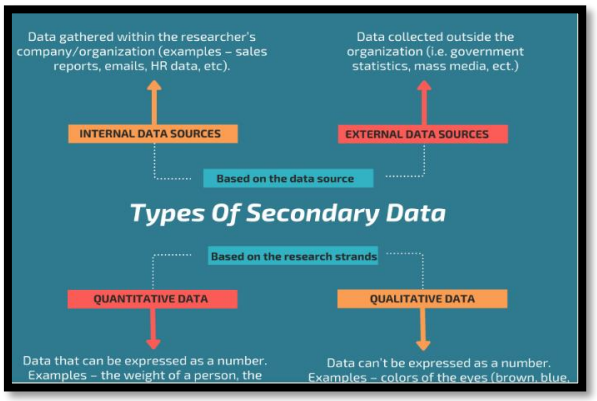

Figure 3.4: Secondary process of data collection

(McGill et al. 2021)

This figure demonstrates the concept of data collection used in this research work with the description of the secondary data. Researchers have collected all information from the "articles, and journals" that are published within the last five years to get the current application approach of ZTA. Moreover, the dataset for visualization has been collected from "Kaggle" for development of the machine learning models to analyze data patterns and the effectiveness of ZTA in increasing network security of APIs. This research work developed the visualization of different data in the considered dataset by identifying their key patterns through the secondary collected data.

3.5 Data analysis

The description of the data analysis is crucial for developing the analytical approach to the key findings of the investigation. Moreover, this provides reliability to the investigation process with the identification of critical findings for the research work. Researchers have applied secondary data all over the research work for the development of critical findings from the data patterns in protecting the data over the APIs network through the application of the ZTA framework. Hence, researchers have created different themes for analyzing the findings through descriptive analysis. This provides a critical evaluation of the overall research work by increasing its acceptability to the target audience. According to Mamabolo and Myres (2020), this analysis process starts with the identification of the pattern of the data involved in the dataset. This provides the concept of analysis approach to the researchers for the construction of different themes. Thus, the development of secondary research focuses on the development of the "thematic approach" for data analysis with the consideration of different hypotheses developed through "conceptual framework" development. All the hypotheses are checked and proved through the investigation and their result is enlisted in the data analysis through different theme creation.

3.6 Evaluation Plan

The development of evaluation planning in the research work is beneficial for developing the evaluation of the research result. Researchers have collected different secondary information and they have evaluated their applications in the real world with the considerations of different security vulnerabilities of APIs. These identifications allowed the researchers to demonstrate the appropriate application of the ZTA in developing the cyber security of APIs. The consequences of different security issues over the ZTA consideration have been analyzed by the researchers for the identification of key security considerations in developing network security with the research work. According to Snyder (2019), the security control considerations have been evaluated by this research work with the development of different valuations of developing architecture of

ZTA for applying in increasing security considerations of APIs. Thus, the result gained from the research work is evaluated with the analysis of its application in increasing cyber security with real-life operations.

3.7 Validation

Validation is a crucial instrument for ensuring the accuracy of the research findings by accessing the most relevant resources for data collection with the development of key findings. This paper is developed with validation with the considerations of only “peer-reviewed” journals for the collection of “secondary data” with the determination of different keywords. Researchers have collected their required secondary resources for reliable resources with the identified aim of those research papers for further development and incorporation of the findings in this investigation (Im *et al.* 2019). The overall data collection techniques are developed with a clear concept of secondary data and its required validation to increase the transparency of the research work. Moreover, researchers have developed an empirical evaluation of all of the findings and considered techniques with proper justification for employing them. This improves the validity of the paper to be adopted and applied by other targeted audiences for increasing the security of APIs.

3.8 Research Ethics

Considerations of “ethical principles” in the research work are crucial for developing transparent and ethical research work with transparent results. Researchers have maintained “confidentiality, and data security” during developing this investigation for the improvement of the transparency of the paper. This research work involved different ethical principles for the implementation of the research work by respecting society and the “human rights” of people. The overall work is developed so that everyone cannot be harmed through the research (Cascio *et al.* 2021). Researchers have developed all required ethnicities for the development of the research by following the anonymity of the other researchers. Moreover, the paper is developed by collecting information from “open access' sources for maintaining the data security and confidentiality of other researchers. Thus, the paper is reliable to be applied in industrial use by society for applying the ZTA framework.

4. Design and Implementation Specifications

4.1 Designing of Zero Trust Architecture for Providing Security

The image below shows the actual implementation of the trust architecture. This architecture is developed with the concept of “never trust, always verify” and by developing three other key trust principles for developing the security of APIs. “Continuous Monitoring and Observability” is another principle incorporated in this network architecture for the development of ZTA applications. “Least Privileges” is another principle that allows only a few bare resources for decreasing network vulnerabilities, along with “Micro Segmentation” of the overall network for creating security in each segment differently.

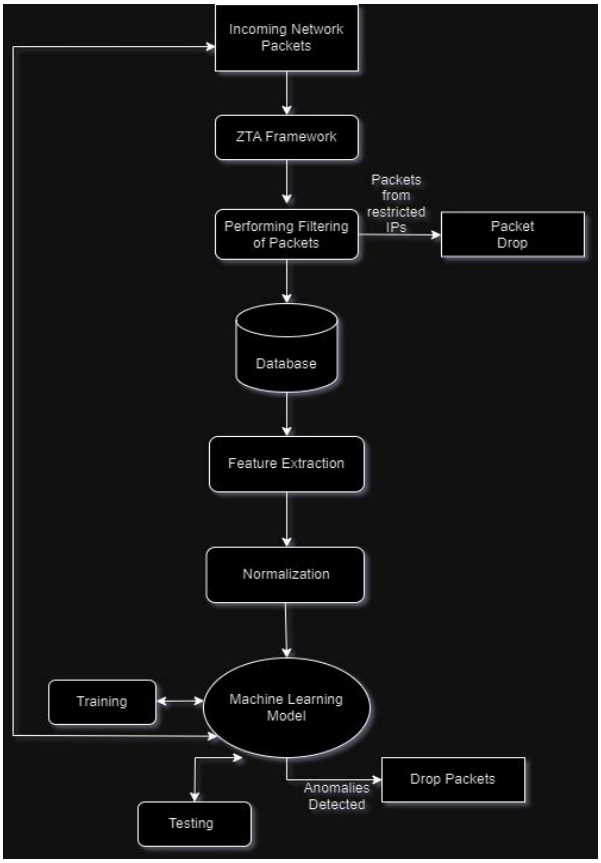

Figure 4.1.1: Proposed flow of Zero Trust architecture framework

Researchers have considered applying the below framework for the development of the ZTA architecture as per the previously considered key principles. This developed "Zero Trust security model" corresponds to the application of APIs in a network that does not trust anyone rather than asking for user authentication. The application of "NIST Zero Trust Architecture (ZTA)" can be beneficial in this case with the consideration of specific solutions to critical security issues within an organization. This architecture includes different components such as "services, sensors and actuators, Internet of Things devices, data repositories, mobile devices" for the development of the final implementation of ZTA in increasing the security of APIs.

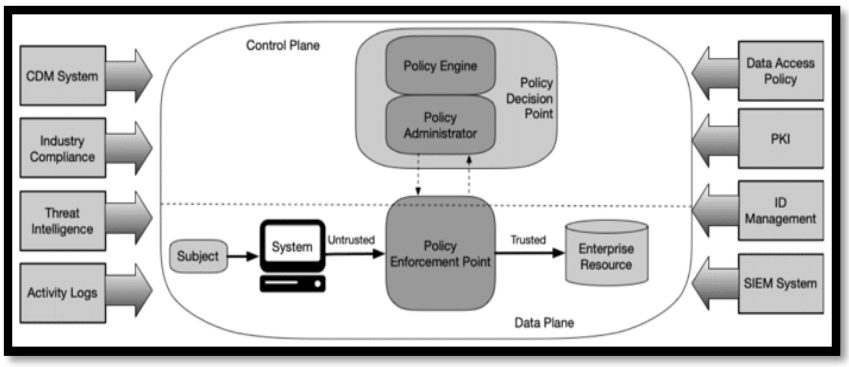

Figure 4.1.2: Developed architecture of the ZTA

(Leahy and Thorpe, 2022)

Researchers have considered applying this framework for the development of the ZTA architecture as per the previously considered key principles. This developed "Zero Trust security model" corresponds to the application of APIs in a network that does not trust anyone rather than asking for user authentication. The application of "NIST Zero Trust Architecture (ZTA)" can be beneficial in this case with the consideration of specific solutions to critical security issues within an organization.

This architecture includes different components such as “services, sensors and actuators, Internet of Things devices, data repositories, mobile devices” for the development of the final implementation of ZTA in increasing the security of APIs.

4.2 Designing of the research work

Analysis of the research design is necessary for the identification of key steps involved in overall research work as per the considered research method. This investigation was developed by following secondary research that considered historical data from secondary resources.

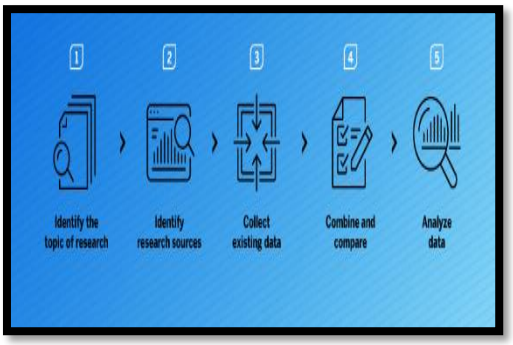

Figure 4.2: Research design
(Rezigalla, 2020)

This figure shows the considered design and steps involved in this research work by following the concept of secondary research. Resources have identified the research area through background analysis for the identification of different resources required for its development. Moreover, by following the research area and its previous researched concept researchers have defined objectives of developing this paper through the identification of data collection processes. After the identification of critical data, researchers analyzed all this collected data through data visualization over **Jupyter Notebook** or **Google colab** with the application of Python Language. Thus, the result has been developed through the data analysis for the identification of key findings of the paper. The overall research work is developed by collecting and analyzing secondary data with the design of secondary research.

4.3 Implementation of the Zero Trust Architecture

The application of Zero Trust Architecture is developed by following its key conceptsand application approach with the considerations of a “data-centric approach”. This is implemented in a network by the identification of different resources and data that is to be protected with the application of ZTA. Thus, consideration of potential threats is necessary for the application of this architecture in increasing the cyber security of APIs. Researchers have developed visibility of the network for the identification of valuable data, assets, applications, and services to be protected (Surantha and Ivan, 2020). The three most crucial factors are developed by architectural development for evaluation of its accessibility and performance. These factors are “users, application, and infrastructure” as these are the most critical factors for the determination of the security effectiveness of the ZTA network.

4.4 Implementation of the research results

The process highlights different kinds of functionalities which must rely on basic fundamentals in each implication in this research methodological chapter.

Overall signified manners are classified in various terms in order to validate the instances by implementing various machine-learning statistics. According to the research works of Campbell (2020), increasingly adopted in various instances are mainly conveying the statistical overviews

in systematic path allocations by which the understandings of each outline are operating. In order to manage the cyber-physical security for factories is generally designing the operative management features in machine learning implications. Addressing each reliable design specified manners in availability and maintainability which can consider each fundamental present in this project. Generalizing each systematic event in comprehending this dataset can use various estimation processes to be proposed by B-profile systems. Overall case studies of Cao (2022) elaborate this accurate profile management is based on abstract behaviors and human interactions which are based on naturalistic benign background. Identifying in base-level interventions some outliers are briefly verified in these kinds of criteria management services. Implemented attacks are addressed in the dataset, in a daily basis format from Tuesday to Friday. This implementation of the dataset elements shows mainly the executed attacks in each area of a week. Some machine learning models are addressed in the process for addressing some valid instances using the training process of the Classifier**.** Entire values engagement values are gaining vast communicative interpretations by using the **XGB Classifier**, **Deep learning models** which can manage the validation process in order to be more adequate in nature assessments.

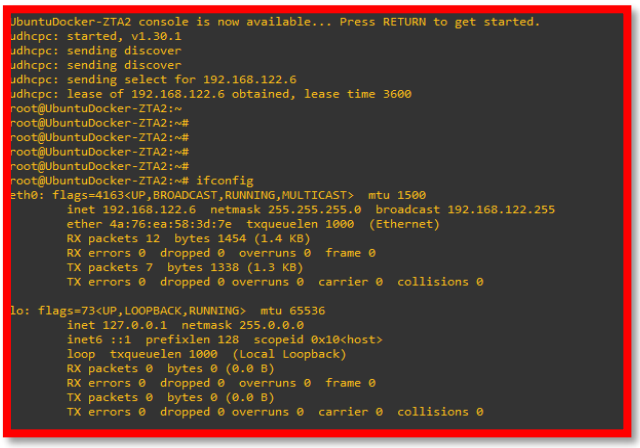

Figure 4.4.1: Docker ZTA configuration

The result displays the network settings for a Docker container with the name "UbuntuDocker-ZTA2." "eth0" for external communication and "lo" for local communication are its two interfaces. The container has an IP address of 192.168.122.6 and a 255.255.255.0 subnet mask. To improve security for unreliable APIs, this configuration is a component of setting up Docker inside a Zero-Trust Architecture (ZTA). Network access is constrained and controlled by isolating the container and assigning it a certain IP address. This enhances the security of the Zero-Trust Architecture by ensuring that communication with vulnerable APIs is restricted inside a secure environment. The figure below shows the infrastructure setup the has been carried out on GNS3 to create a kind of ZTA framework.

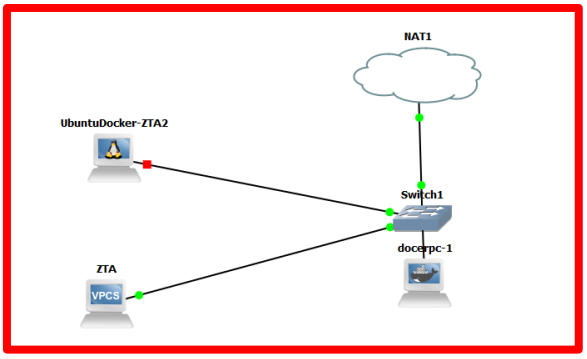

Figure 4.4.2: Infrastructure Setup.

4.5 Feature Selection

This method is one of the activities in the machine learning which helps to boost the efficiency of a predictive model via identifying a particular group of attributes from a huge dataset here, also the chosen dataset is having huge entries and we are selecting only a subset called DOS and benign for the implementation. By carrying out this methodology it improves the accuracy of the proposed model and helps to provide an easy way for the interpretation of the results.

In this implementation features are selected using a method called **co-relation-based feature selection** which is also known as CFS method. In this particular method that identifies the attributes according to the extent to which they associated with the target value and other attributes. Once after segregation of the attributes from the target dataset only the relevance attributes are used for the model training. Below snapshot describes the absolute attributes that are used for training the proposed ML model.

	Unnamed: 0	Flow ID	Src IP	Src Port	Dst IP	Dst Port	Protocol	Timestamp	Flow Duration	Tot Fwd Pkts	Tot Bwd Pkts	TotLen Fwd Pkts	TotLen Bwd Pkts
0	624	192.168.4.118-203.73.24.75-4504-80-6	192.168.4.118	4504	203.73.24.75	80	6	12/06/2010 08:34:32 AM	3974862	29	44	86.0	59811.0
1	625	192.168.4.118-203.73.24.75-4504-80-6	192.168.4.118	4504	203.73.24.75	80	6	12/06/2010 08:34:36 AM	63	1	1	0.0	0.0
2	626	192.168.4.118-203.73.24.75-4505-80-6	192.168.4.118	4505	203.73.24.75	80	6	12/06/2010 08:34:36 AM	476078	2	6	86.0	3037.0
3	627	192.168.4.118-203.73.24.75-4505-80-6	192.168.4.118	4505	203.73.24.75	80	6	12/06/2010 08:34:37 AM	151	2	1	0.0	0.0
4	628	192.168.4.118-203.73.24.75-4506-80-6	192.168.4.118	4506	203.73.24.75	80	6	12/06/2010 08:34:37 AM	472507	2	5	73.0	1050.0

Figure 4.5: Output of Feature selection.

4.6 Training and Testing of ML model:

Once the selected dataset has been cleaned and pre-processed, another ML operation called Training and testing has been performed. In this scenario the selected dataset has been divided into two main parts training and testing. As the whole dataset is being divided into two main parts in which 60% of the data is being used for training the model and the remaining 40% is being used for testing the model. The entire working dataset has been selected from Kaggle repository. There are no recent datasets found in public domain that can be used for DDOS and other kinds of attacks. The proposed model is being trained basically based on ML models like Random Forest Classifier and XGBoost Classifier and with some Deep learning models and recommending the best trained model based on the accuracy score of the model. Lastly, the testing data is utilized for determining the trained model accuracy and f1 score.

4.7 Metrics for Model Testing:

True Positive rate: It is a metric that evaluates the proportion of properly identified actual findings. Mathematically can be identified as follows:

$$\text{True Positive rate} = \frac{tp}{fn+tp}$$

False Positive rate: It is the amount of illustrations among all samples that were not a part of class X that were classified as class benign, expressed mathematically as:

$$\text{False Positive rate} = \frac{fp}{fp+tn}$$

Precision: It is the amount of the classifier's actual positives to all the experiment's positive outcomes, which can be expressed mathematically as shown below:

$$\text{Precision} = \frac{tp}{tp+fp}$$

Recall: It is practically defined as ratio of actual positives to all possible occurrences and mathematically represented as shown below:

$$\text{Recall} = \frac{tp}{tp+fn}$$

Accuracy Score: It is one of the most typical and widely utilized metrics to assess a model and collect output results. Employing the Python sklearn library, this accuracy score is determined. Accuracy score produced by this approach fluctuates depending on the ML algorithms employed and mathematically expressed as shown below:

$$\text{Accuracy Score} = \frac{(Tp+Tn)}{(Tp+Tn+Fp+Fn)}$$

Classification Report: The classification report is an additional parameter metric utilized for the assessment of output. The study indicates it is an accurate resource for the machine learning technique's analysis. The most widely utilized Python library, sklearn, is the one that is used in the categorization process. The effective execution of computations is made possible by making use of this package. In accordance with the algorithm employed, the classification summary utilizes other strategies. A range of factors, comprising Precision, Recall, F1 score, Support, etc., are covered in the classification analysis. This parameter's calculation should result in a ratio of accurately estimated positive samples to all anticipated samples that are positive.

5. Evaluation

In this particular section, the findings of applying the proposed machine learning models will be discussed. At the same time in this section the different metrics discussed in the above section will be evaluated between the selected ML models. The ML models are XGBOOST, Random Forest Classifier and Deep learning models like CNN, LSTM, BILSTM and BILSTM with attention mechanism. The specified ML models have been implemented and accuracy of each model in detecting the network anomalies within ZTA architecture are summarized in the below table. The results show that BILSTM with attention mechanism has achieved highest accuracy and it is about 99%, while other models like BILSTM achieved 95.82%, LSTM achieved 93.16%, CNN achieved 88.43%, and other ML models like XGBOOST achieved 95.70%, Random Forest Classifier achieved 94.21%. The below table gives detailed information of each developed ML model.

ML Model	Accuracy	Precision	Recall	F1-Score
Random Forest Classifier	94.21%	97%	91%	94%
XGBOOST Classifier	95.70%	92%	100%	96%
CNN	88.43%	94%	82%	88%
LSTM	93.16%	98%	88%	93%
BILSTM	95.82%	97%	95%	96%
BILSTM with Attention Mechanism	99.25%	100%	99%	99%

5.1 Scenario 1: Random Forest Classifier

The accuracy of Random Forest Classifier is 94.21%. Below snapshot explains the performance evaluation of the model with the help of Confusion Matrix.

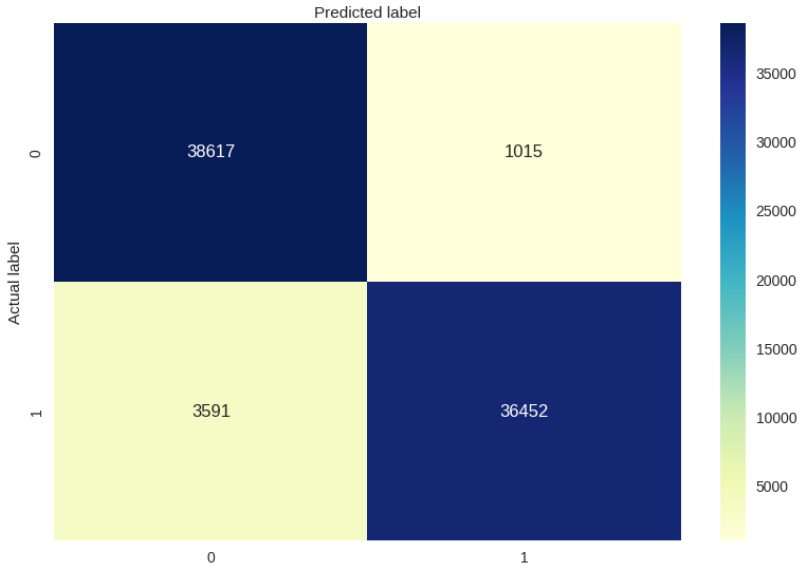

Figure 5.1: Confusion Matrix of Random Forest Classifier.

Below table summarizes the confusion matrix shown above.

Class Name	Number of Correct Classified Data Samples	Number of Incorrectly Classified Data Samples
DDOS	36452	3591
Beningn	38617	1015

5.2 Scenario 2: XGBOOST Classifier

The accuracy of XGBOOST Classifier is 95.70%. Below snapshot explains the performance evaluation of the model with the help of Confusion Matrix.

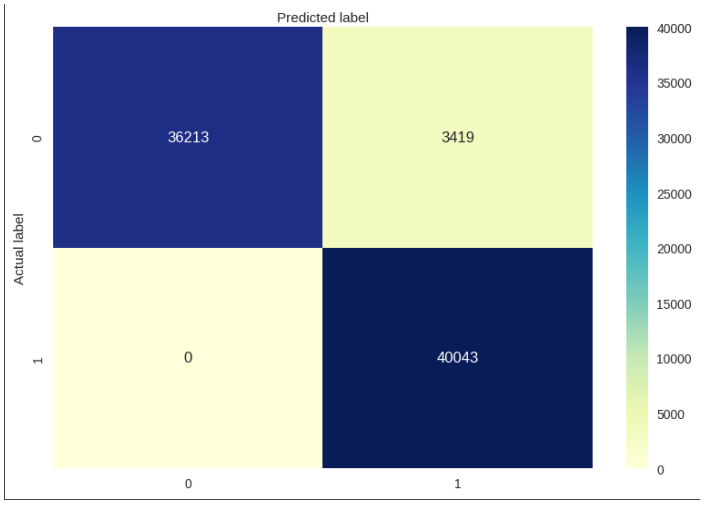

Figure 5.2: Confusion Matrix of XGBOOST Classifier.

Below table summarizes the confusion matrix shown above.

Class Name	Number of Correct Classified Data Samples	Number of Incorrectly Classified Data Samples
DDOS	40043	0
Beningn	36213	3419

5.3 Scenario 3: Deep Learning Model CNN

The accuracy of CNN model is 88.43%. Below snapshot explains the performance evaluation of the model with the help of Confusion Matrix.

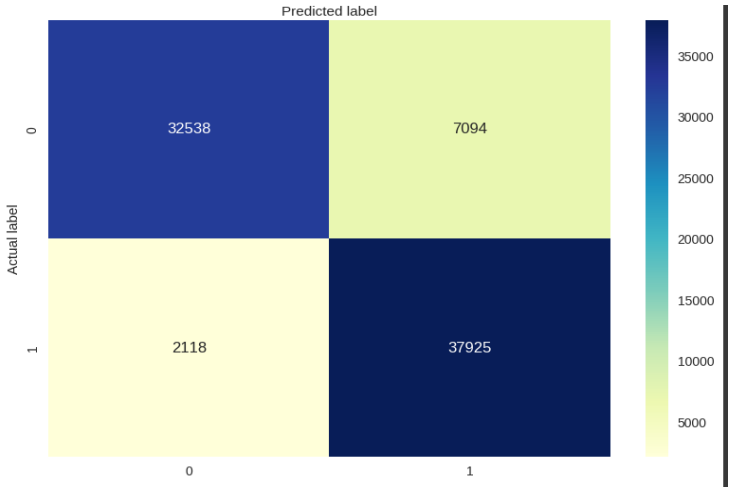

Figure 5.3: Confusion Matrix of CNN.

Below table summarizes the confusion matrix shown above.

Class Name	Number of Correct Classified Data Samples	Number of Incorrectly Classified Data Samples
DDOS	37925	2118
Beningn	32538	7094

5.4 Scenario 4: Deep Learning Model LSTM

The accuracy of LSTM model is 93.16%. Below snapshot explains the performance evaluation of the model with the help of Confusion Matrix.

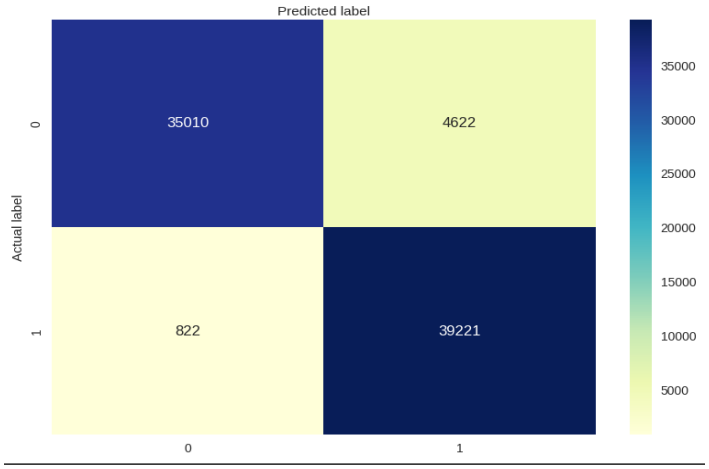

Figure 5.4: Confusion Matrix of LSTM.

Below table summarizes the confusion matrix shown above.

Class Name	Number of Correct Classified Data Samples	Number of Incorrectly Classified Data Samples
DDOS	39221	822
Beningn	35010	4622

5.5 Scenario 5: Deep Learning Model BILSTM

The accuracy of BILSTM model is 95.82%. Below snapshot explains the performance evaluation of the model with the help of Confusion Matrix.

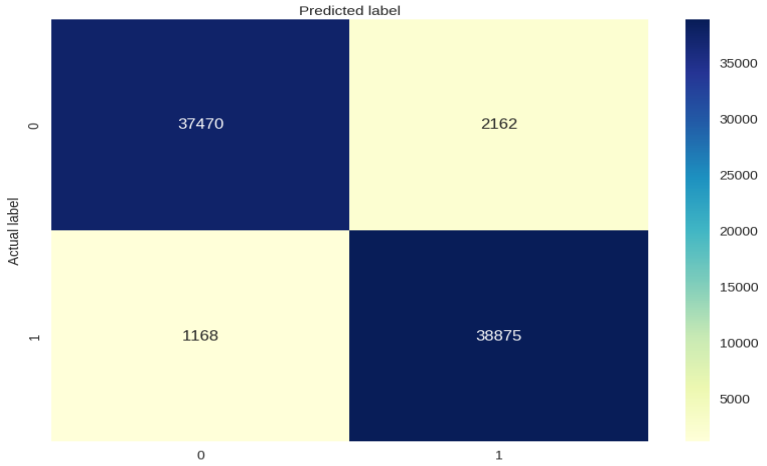

Figure 5.5: Confusion Matrix of BILSTM.

Below table summarizes the confusion matrix shown above.

Class Name	Number of Correct Classified Data Samples	Number of Incorrectly Classified Data Samples
DDOS	39221	822
Beningn	35010	4622

5.6 Scenario 6: Deep Learning Model BILSTM with attention mechanism

The accuracy of BILSTM with attention mechanism model is 99.25%. Below snapshot explains the performance evaluation of the model with the help of Confusion Matrix.

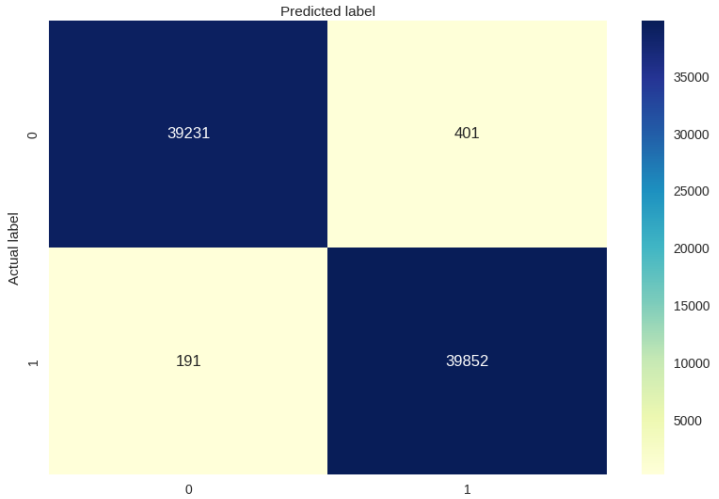

Figure 5.6: Confusion Matrix of BILSTM with attention mechanism.

Below table summarizes the confusion matrix shown above.

Class Name	Number of Correct Classified Data Samples	Number of Incorrectly Classified Data Samples
DDOS	39852	191
Beningn	39231	401

6. Conclusion

APIs can provide threats to security; in that action the main concept of this assessment is to improve the security system of APIs with the help of the *Zero Trust Strategy* through a machine learning approach. The entire implementation of ML models to detect the network anomalies was done using the dataset obtained from the Kaggle repository. The selected dataset was used for models like Random Forest Classifier, XGBOOST Classifier and for Deep learning models like CNN, LSTM, BILSTM, BILSTM with Attention mechanism. And each model produced accuracy which were varying in percentage. Taking accuracy into consideration BILSTM with attention mechanism achieved the highest when compared to other models. Hence, my work focuses on recommending the best model that could give the highest result of security within Zero trust architecture. The proposed work also aims on the challenges associated with respective attacks. By taking different types of attacks which can cause major threat to APIs and computing system, developing a secured system proved to maintain the overall security of the system. However, **t**he techniques presently in operation are still not specifically designed to stop the harmful attacks that are being conducted. Hence, the goal of the study revolved around investigating the attacks and establishing a co-relation between model performances and design specifications.

References

Ameer, S., Gupta, M., Bhatt, S. and Sandhu, R., 2022, June. Bluesky: Towards convergence of zero trust principles and score-based authorization for iot enabled smart systems. In Proceedings of the 27th ACM on Symposium on Access Control Models and Technologies (pp. 235-244).

Chen, B., Qiao, S., Zhao, J., Liu, D., Shi, X., Lyu, M., Chen, H., Lu, H. and Zhai, Y., 2020. A security awareness and protection system for 5G smart healthcare based on zero-trust architecture. IEEE Internet of Things Journal, 8(13), pp.10248-10263.

Csikor, L., Ramachandran, S. and Lakshminarayanan, A., 2022, December. ZeroDNS: Towards Better Zero Trust Security using DNS. In Proceedings of the 38th Annual Computer Security Applications Conference (pp. 699-713).

He, Y., Huang, D., Chen, L., Ni, Y. and Ma, X., 2022. A survey on zero trust architecture: Challenges and future trends. Wireless Communications and Mobile Computing, 2022

Horne, D. and Nair, S., 2021. Introducing zero trust by design: Principles and practice beyondthe zero trust hype. In Advances in Security, Networks, and Internet of Things (pp. 512-525). Springer.

Leahy, D. and Thorpe, C., 2022, March. Zero trust container architecture (ztca): A framework for applying zero trust principals to docker containers. In International Conference on Cyber Warfare and Security (Vol. 17, No. 1, pp. 111-120).

Liu, Y., Hao, X., Ren, W., Xiong, R., Zhu, T., Choo, K.K.R. and Min, G., 2022. A blockchain-based decentralized, fair and authenticated information sharing scheme in zero trust internet-of-things. IEEE Transactions on Computers, 72(2), pp.501-512.

Mehraj, S. and Banday, M.T., 2020, January. Establishing a zero trust strategy in cloud computing environment. In 2020 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-6). IEEE.

Phiayura, P. and Teerakanok, S., 2023. A Comprehensive Framework for Migrating to Zero Trust Architecture. IEEE Access, 11, pp.19487-19511.

Syed, N.F., Shah, S.W., Shaghaghi, A., Anwar, A., Baig, Z. and Doss, R., 2022. Zero trust architecture (zta): A comprehensive survey. IEEE Access, 10, pp.57143-57179.

Teerakanok, S., Uehara, T. and Inomata, A., 2021. Migrating to zero trust architecture: Reviews and challenges. Security and Communication Networks, 2021, pp.1-10.

Newman, M. and Gough, D., 2020. Systematic reviews in educational research: Methodology, perspectives and application. Systematic reviews in educational research: Methodology, perspectives and application, pp.3-22.

Alharahsheh, H.H. and Pius, A., 2020. A review of key paradigms: Positivism VS interpretivism. Global Academic Journal of Humanities and Social Sciences, 2(3), pp.39-43.

Rashid, Y., Rashid, A., Warraich, M.A., Sabir, S.S. and Waseem, A., 2019. Case study method: A step-by-step guide for business researchers. International journal of qualitative methods, 18, p.1609406919862424.

McGill, E., Er, V., Penney, T., Egan, M., White, M., Meier, P., Whitehead, M., Lock, K., de Cuevas, R.A., Smith, R. and Savona, N., 2021. Evaluation of public health interventions from a complex systems perspective: a research methods review. Social Science & Medicine, 272, p.113697.

Mamabolo, A. and Myres, K., 2020. A systematic literature review of skills required in the different phases of the entrepreneurial process. Small Enterprise Research, 27(1), pp.39-63.

Snyder, H., 2019. Literature review as a research methodology: An overview and guidelines. Journal of business research, 104, pp.333-339.

Im, G.H., Shin, D. and Cheng, L., 2019. Critical review of validation models and practices in language testing: their limitations and future directions for validation research. Language Testing in Asia, 9(1), pp.1-26.

Cascio, M.A., Weiss, J.A. and Racine, E., 2021. Making autism research inclusive by attending to intersectionality: A review of the research ethics literature. Review Journal of Autism and Developmental Disorders, 8, pp.22-36.

Rezigalla, A.A., 2020. Observational study designs: Synopsis for selecting an appropriate study design. Cureus, 12(1).

Surantha, N. and Ivan, F., 2020. Secure kubernetes networking design based on zero trust model: A case study of a financial service enterprise in indonesia. In Innovative Mobile and Internet Services in Ubiquitous Computing: Proceedings of the 13th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS-2019) (pp. 348-361). Springer International Publishing.

Campbell, M., 2020. Beyond zero trust: Trust is a vulnerability. Computer, 53(10), pp.110-113.
Cao, Y., Pokhrel, S.R., ZHU, Y., Ram Mohan Doss, R. and Li, G., 2022. Automation and Orchestration of Zero Trust Architecture: Potential Solutions and Challenges.

Cao, Y., Pokhrel, S.R., ZHU, Y., Ram Mohan Doss, R. and Li, G., 2022. Automation and Orchestration of Zero Trust Architecture: Potential Solutions and Challenges.