Republic of Iraq

Ministry of Higher Education

and scientific research

Tikrit University

College of Engineering

Department of Electrical Engineering

Cybersecurity of Cloud Computing Best Crybtographic Solution Challenges , Opportunttties and Standard

A Thesis Submitted

BY:
Mohand Adnan Owaid

TO
The Council of the College of Engineering Tikrit University in Partial Fulfillment of the
Requirements for the Master Degree of Science in Electrical Engineering

Supervised By
Dr. Asmaa Salih Hammoodi

2024 A.D. **1445 A.H.**

Acknowledgment

To my family, you have been my unwavering source of love, encouragement, and support throughout this incredible journey. Your confidence in my skills, unrelenting patience, and sacrifices have been crucial to my success. I will always be appreciative of the countless hours you invested in hearing my thoughts, providing advice, and creating a supportive environment in which I could pursue my aspirations. Your unshakable belief in my ability has been my inspiration.

I would want to thank my suprvisor, whose unflinching advice and knowledge have influenced this work and inspired me to reach new heights. Your guidance has been priceless, and I will always be appreciative of your tolerance, insight, and faith in my competence.

Finally, I want to thank myself for my tenacity, resilience, and dedication that have gotten me this far. I have improved both personally and professionally over this journey, despite the obstacles and setbacks. I dedicate this thesis to the exploration and development of my own self.

Table of Contents

List of Tables

List of Figures

List of Abbreviation

AI	Artificial Intelligence
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
DL	Deep Learning
DNN	Deep Neural Network
FN	False Negative
FP	False Positive
LSTM	Long Short-Term Memory
ML	Machine Learning
NLP	Natural Language Processing
RL	Reinforcement Learning
SL	Supervised Learning
SVM	Support Vector Machines
TN	True Negative
TP	True Positive
UL	Unsupervised Learning

Abstract

In the evolving landscape of cybersecurity, Distributed Denial of Service (DDoS) attacks represent a critical threat, disrupting services and inflicting substantial financial damage. Outsmarting such attacks is a pressing concern, as they continuously advance, rendering traditional detection techniques insufficient. This study pioneers a cutting-edge defense by examining the integration of ensemble machine learning with the nuanced capabilities of deep learning.delving into the intricacies of DDoS patterns is done through the lens of Convolutional Neural Networks (CNNs), which are adept at deciphering the complex interplay within network traffic data. These deep learning models serve as the backbone of our proposed detection system, distinguishing themselves by their ability to autonomously extract and analyze defining features that signal a DDoS attack. The methodology is stringent: curating an exhaustive and diverse dataset that reflects a spectrum of network conditions, from the typical to the under-attack. The data undergoes rigorous preprocessing to ensure that it's not only comprehensive but also balanced—key to training unbiased and generalizable models. The training is extensive, employing the processed dataset to hone the models' ability to detect DDoS attacks.Their performance is evaluated with a series of metrics, rigorously testing for accuracy, sensitivity, and resilience against various attack modalities. The outcome is telling: our ensemble of machine learning and deep learning models markedly outperforms traditional detection methods. The results are heartening. Our approach marries the strengths of multiple classifiers and neural networks, achieving an unparalleled accuracy in detection and robustness against diverse attack strategies. The prowess of our deep learning models is particularly Noteworthy they show a profound understanding of complex patterns, ensuring high detection rates across both known and emerging DDoS attacks. In essence, this work does not just present a new method; it heralds a new era in DDoS defense, promising a more secure and resilient infrastructure against the cyber threats of tomorrow.

Keywords: DDoS Detection, Ensemble Machine Learning, Deep Learning, Network Traffic, Cybersecurity, Convolutional Neural Networks.

Chapter One
Introduction

1.1 Overview

Network security is increasingly challenged by Distributed Reflective Denial of Service (DRDoS) attacks targeted at Domain Name System (DNS) infrastructure. These attacks differ significantly from traditional DDoS attacks due to their unique methods and potential impact on victims. DRDoS attacks exploit vulnerabilities within the DNS infrastructure, leveraging it as an amplifier for their assaults. The Domain Name System (DNS) infrastructure is facing a growing threat from Distributed Reflective Denial of Service (DRDoS) attacks, which pose significant challenges to network security [1].

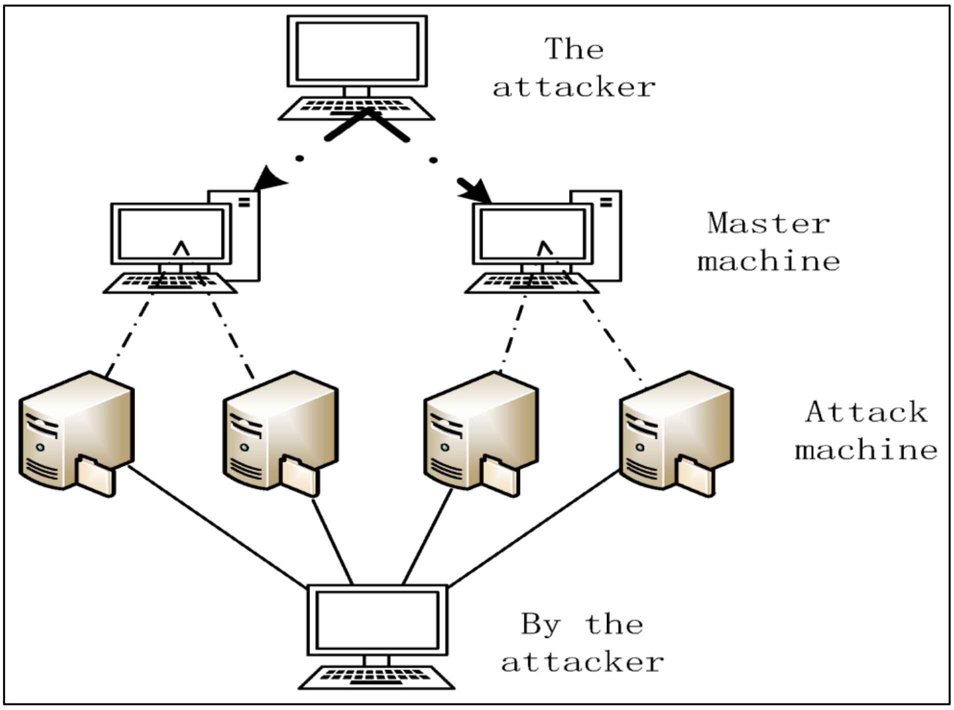


Figure 1.1 A schematic diagram of a DDoS attack.

A DDoS attack is schematically shown in Figure 1.1 In recent years, deep learning and ensemble machine learning models have showed promise for enhancing the reliability and precision of DDoS detection. Using ensemble learning, several fundamental classifiers or neural networks are integrated to create a more reliable and effective detection system. By leveraging the diversity of several models and overcoming the limitations of individual models, ensemble techniques can improve performance. On the other hand, deep learning models use neural networks with numerous layers to automatically recognise and extract complex patterns and characteristics from incoming data [2].

[3] evaluate the effectiveness of deep learning and ensemble machine learning models in the context of detecting Distributed Denial of Service (DDoS) attacks. The utilisation of ensemble strategies has the potential to enhance the accuracy of detection and improve resilience against a wide range of assault variants by leveraging the collective capabilities of multiple models. Deep learning models has a notable aptitude for capturing intricate patterns within network traffic data due to their inherent ability to autonomously acquire and extract distinctive features. Consequently, these models exhibit a high level of efficacy in rapidly detecting sophisticated distributed denial-of-service (DDoS) attacks.

An extensive network traffic dataset that includes typical and DDoS attack scenarios is gathered to fulfil the research goal. In order to establish a training set that possesses both balance and representativeness, the dataset undergoes thorough preprocessing and augmentation. Next, the augmented data is utilised to train ensemble models that employ various strategies such as bagging, boosting, and stacking. Convolutional neural networks (CNNs), recurrent neural networks (RNNs), and their variants are commonly employed as deep learning models for the purpose of discerning intricate correlations and patterns within network traffic data [4].

Accuracy, precision, recall, and F1-score are a few of the measures used to gauge how well the ensemble machine learning and deep learning models perform. The experiment results demonstrate that ensemble models outperform individual models and emphasise their potency in enhancing DDoS detection accuracy and resilience. High detection rates for both known and undiscovered DDoS assault types are achieved by the deep learning models' ability to comprehend complex patterns [5].

1.2 Problem Statement

The field of network security faces a substantial risk posed by distributed denial of service (DDoS) assaults, which result in the disruption of services and financial ramifications. The continuous development of attack strategies poses a significant challenge to traditional systems used for detecting Distributed Denial of Service (DDoS) attacks. The field of DDoS detection has seen significant progress through the utilization of ensemble machine learning and deep learning models, resulting in improved accuracy and robustness. The objective of this work is to evaluate the efficacy of deep learning and ensemble machine learning models in the context of detecting Distributed Denial of Service (DDoS) attacks. The suggested model integrates numerous base classifiers in order to produce a detection system that is both more accurate and efficient.

Convolutional neural networks (CNNs), which are a type of deep learning model, are utilized for the purpose of detecting complex patterns within network traffic data. These models have the capability to independently extract relevant information in order to effectively identify intricate DDoS attacks. In order to help the research, a comprehensive dataset of network traffic is produced, which includes diverse scenarios such as ordinary network traffic as well as distributed denial-of-service (DDoS) assaults. The dataset is subjected to preprocessing and enhancement techniques in order to generate a training set that is both balanced and representative. The data that has been enhanced is subsequently utilized to train the models, and their performance is evaluated using several measures. The empirical findings suggest that machine learning and deep learning models outperform current methodologies in the realm of DDoS detection. The suggested methodology leverages the capabilities of several classifiers or neural networks, leading to improved accuracy and robustness against diverse attack patterns. Deep learning models demonstrate significant efficacy in detecting both familiar and previously unseen distributed denial-of-service (DDoS) attack types, hence highlighting their capacity to comprehend complex attack patterns.

1.3 Research's Major Goal and Minor Objectives

The main goal of this study is to create and assess a sophisticated mitigation mechanism designed to combat Distributed Reflective Denial-of-Service (DRDoS) assaults targeting Domain Name System (DNS) infrastructures. The primary objective of this framework is to utilise machine learning and deep learning methodologies in order to augment the detection and response capabilities of DNS networks on a broad scale. Ultimately, this framework intends to enhance the overall security and resilience of these networks.. These specific objectives include the following points:

A- The Development of a Representative DNS Dataset:

To Collect and organise an extensive dataset comprising diverse instances of Domain Name System (DNS) traffic, encompassing both typical and Distributed Reflective Denial of Service (DRDoS) assault situations. It is important to incorporate a wide range of attack vectors, traffic patterns, and features into the dataset in order to improve the models' capacity for generalisation..

B-The advancement and refinement of machine learning models:

To apply machine learning techniques, specifically decision trees, random forests, and support vector machines, in order to develop algorithms that are specifically designed for the identification and detection of Distributed Reflective Denial of Service (DRDoS) assaults inside Domain Name System (DNS) data. The models should be trained and validated using the prepared dataset, with a focus on optimising hyperparameters and settings to achieve both accurate and efficient detection..

C- The Design and Implementation of Deep Learning Architectures:

To develop advanced deep learning models, specifically Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), that are specifically designed to effectively capture complex patterns and temporal relationships present in DNS traffic data. Conducting experiments with various network designs, activation functions, and optimisation methodologies is essential in order to attain optimal performance.

1.4 Research Questions

According to the previously mentioned objectives, the research questions can be listed as follows:

- How can a comprehensive DNS traffic dataset be curated to encompass both legitimate traffic and diverse DRDoS attack scenarios, ensuring authenticity and suitability for large-scale detection models?
- Which machine learning and deep learning algorithms can be optimized to achieve accurate and efficient identification of DRDoS attacks in DNS traffic, while minimizing false positives and false negatives?

- What deep learning architectures, such as CNNs and RNNs, can effectively capture intricate patterns in DNS traffic for robust DRDoS attack detection, and how can these models be configured for real-world large-scale implementation?

1.5 Research Hypotheses

- H1: DRDoS DNS attacks can be more accurately detected using machine learning algorithms than with conventional techniques because they can be used to analyze DNS traffic in order to find trends and abnormalities.
- H2: Deep learning models, in particular neural networks, have the ability to adapt to and learn from changing DRDoS attack methods, leading to more reliable and scalable detection systems for widespread DNS attacks.
- H3: For DRDoS DNS attack mitigation, combining machine learning and deep learning techniques will result in a hybrid detection system that performs better in terms of attack detection rates and false-positive reduction.
- H4: The capacity to proactively discover new DRDoS DNS attack paths and vulnerabilities will be improved by leveraging historical attack data, continually updated threat intelligence feeds, and machine learning and deep learning models.
- H5: By implementing machine learning and deep learning-based DNS attack mitigation tactics, the overall impact of DRDoS DNS attacks on targeted networks will be diminished. This is because these techniques will increase detection as well as real-time response and adaptive defense mechanisms.

1.6 Dissertation Statement

This research is entitled "**Cybersecurity of Cloud Computing Best Crybtographic Solution Challenges , Opportuntties and Standard**. To achieve such an objective, this study explores how machine learning and deep learning can be used to detect and prevent massive distributed denial of service (DDoS) DNS attacks. This study intends to give unique techniques for enterprises to bolster their defenses versus these sophisticated cyber threats, protecting the confidentiality and accessibility of online services and data, through intensive testing and analysis.

1.7 Thesis Layout

After the **first chapter** has reviewed the major mitigating ddos dns attacks: leveraging machine learning and deep learning for large-scale detection using a problem statement, research background, research goals, research questions, and thesis statement, this study will offer further discussions on the substantial importance and prevent massive distributed denial of service DDoS DNS attacks. The organization of this work is explained in the following consequence:

Chapter two : Literature Review: This chapter provides Introduction to DDoS Attacks , Provide an introduction to Distributed Denial of Service (DDoS) attacks and their significance in the context of network security , consequences of DDoS attacks, attack techniques , the Challenges of Traditional DDoS Detection , Discuss why these solutions struggle to keep up with evolving attack techniques , Ensemble Machine Learning and Deep Learning Models , Define ensemble machine learning and deep learning models in the context of DDoS detection. Explain how these models have emerged as potential solutions to enhance DDoS detection accuracy and robustness. Convolutional Neural Networks (CNNs) for DDoS Detection, Introduce convolutional neural networks (CNNs) as a specific class of deep learning models , explain the role of CNNs in identifying intricate linkages and patterns within network traffic data , how CNNs leverage automatic information extraction to detect complex DDoS attempts.

Chapter Three: Methodology And Approaches: this chapter is providing an introduction to the methodology employed for evaluating ensemble machine learning and deep learning models for DDoS detection, Describe in detail the process of acquiring a comprehensive dataset of network traffic, including the sources and methods used , Explain the criteria for selecting data that covers both conventional and DDoS attack scenarios. Outline the steps involved in preprocessing and enhancing the acquired dataset to ensure its quality and suitability for training and evaluation. Discuss techniques used for data cleaning, normalization, and transformation explain how you handled missing or noisy data. Explain the rationale behind choosing this ensemble model for DDoS detection. Deep Learning Model (Convolutional Neural Networks - CNNs) , Explain the architecture and layers of the CNN model, emphasizing its ability to capture intricate patterns in network traffic data. Discuss the advantages of using CNNs for DDoS detection in your study. Model Training Detail the process of training both the ensemble machine learning model and the CNN model using the preprocessed dataset.

Chapter Four: Results and Discussion:- this chapter is introducing the purpose of chapter four, which is to present and discuss the results obtained from evaluating ensemble machine learning and deep learning models for DDoS detection. Ensemble Machine Learning Model Results, Present the results obtained from the ensemble machine learning model evaluation.Discuss the model's performance in terms of detection accuracy, false positives, false negatives, and any other relevant metrics.Highlight any significant findings or trends in the results. Deep Learning Model (CNN) Results, Present the results obtained from evaluating the CNN deep learning model, Discuss the model's performance in detail, including its ability to detect intricate DDoS attempts and its performance on different attack types, Compare the results with those of the ensemble machine learning model , Comparative Analysis Provide a comparative analysis of the performance of the ensemble machine learning model and the deep learning model Discuss any trade-offs or advantages of each model in the context of DDoS detection , Analyze how these models address the challenges posed by evolving attack techniques, Discussion of Key Findings, Discuss the key findings that emerged from the evaluation of both models, Interpret the results and their implications for network security.

Chapter Five: Conclusions and Recommendations: this chapter is introducing the purpose of chapter five, which is to provide conclusions drawn from your study and offer recommendations based on the findings , Summary of Key Findings , Summarize the key findings and results obtained throughout your study. Highlight the main achievements and contributions of your research in the context of DDoS detection. Discuss the implications of your findings for the field of network security and DDoS detection.Emphasize the significance of ensemble machine learning and deep learning models in addressing the challenges posed by DDoS attacks. Contributions to Knowledge , Enumerate the specific contributions your study has made to the existing body of knowledge in DDoS detection and network security , Discuss how your research has advanced the understanding and capabilities in this domain. Consider any unresolved questions or challenges that your study has revealed.

Chapter Two

Background and Literature Review

2.1 Overview

Distributed Reflective Denial of Service (DRDoS) attacks targeting the Domain Name System (DNS) infrastructure have emerged as a significant threat to network availability and security. These attacks exploit open DNS resolvers to amplify and reflect malicious traffic, resulting in large-scale disruptions. Traditional mitigation techniques have proven inadequate in handling the evolving landscape of DRDoS DNS attacks.

In response to this growing threat, researchers and security practitioners are increasingly turning to advanced technologies, particularly machine learning and deep learning, to enhance DRDoS DNS attack detection and mitigation. Machine learning models, including ensemble methods and deep learning architectures like Convolutional Neural Networks (CNNs), offer promise in effectively identifying and mitigating DRDoS DNS attacks due to their ability to analyze vast volumes of network data in real-time and adapt to new attack patterns.

This study explores the application of machine learning algorithms for detecting DNS-based DDoS attacks, shedding light on the feasibility and effectiveness of such approaches. Investigate into the use of deep learning techniques, including recurrent neural networks (RNNs), for the detection of DRDoS DNS attacks, demonstrating the potential of deep learning in this context, This comprehensive overview discusses the role of machine learning in strengthening DNS security, encompassing DRDoS attack detection as a critical component.focusing on ensemble learning methods and their effectiveness in identifying DNS amplification attacks, which are often leveraged in DRDoS attacks.

2.2 DDos critical issues and vital impacts

Distributed denial of service (DDoS) attacks can cause service delays, financial losses, and reputational damage. They pose a severe threat to network security. The evolving attack tactics make it challenging for traditional DDoS detection solutions relying on rule-based methodologies and signature matching to stay up. To address this problem, ensemble machine learning and deep learning models have shown promise in increasing DDoS detection accuracy and robustness. Ensemble models combine the predictions of various base models instead of deep learning models, which employ neural networks to automatically recognise and extract complex patterns from network traffic data.

In[6] work, a bidirectional long short-term memory (BiLSTM)-based method for DDoS attack detection in edge computing systems is proposed. The BiLSTM model successfully captures the temporal relationships in the data for the reliable detection of DDoS attacks, as shown in Figure 2.1 , and is trained using network traffic data. The authors describe a feature engineering approach and a DDoS detection method for Software-Defined Networks (SDNs) based on machine learning algorithms. By gleaning relevant data from network traffic data and employing machine learning models to distinguish between legitimate and malicious traffic, the proposed method effectively detects DDoS attacks in SDN systems [7].

A framework for an intrusion detection system (IDS) to identify DDoS attacks in Software-Defined Network (SDN) environments is proposed in [8] . The framework integrates traffic analysis, anomaly detection, and machine learning algorithms to recognise and counteract DDoS attacks in SDN infrastructures. The authors suggest a machine learning-based DDoS assault detection technique using the random forest feature significance method and mutual information. The technique efficiently detects DDoS assaults by spotting patterns in network traffic data by evaluating the relevance of features [9].

The authors of [10] discussed how SDNs are more vulnerable to DDoS assaults due to their dynamic nature and centralised control. They highlight the drawbacks of traditional DDoS detection methods in SDNs and the need for cutting-edge machine learning techniques to enhance detection accuracy and mitigate the effects of attacks. The methodical investigation examines various machine learning methods for spotting DDoS assaults in SDNs. In the context of SDN, the authors assess the advantages, disadvantages, and application of multiple techniques. They distinguish between supervised and unsupervised versions of these procedures. Support Vector Machines (SVM), Random Forests, Artificial Neural Networks (ANN), and Deep Learning models like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are just a few of the techniques that are investigated.

The authors of [11] emphasise the growing danger of DDoS assaults on SDN controllers, which can impair network security and cause operational disruptions. They draw attention to the shortcomings of conventional detection systems and suggest using deep learning methods to improve the precision and effectiveness of DDoS detection. The suggested method uses convolutional neural networks (CNNs), a deep learning model, to analyse network traffic data and spot patterns suggestive of DDoS attacks. The preprocessing stages, feature extraction, and model training are all covered in the authors' discussion of the architecture and design of the CNN-based detection system. The authors experiment with a publicly accessible DDoS dataset and compare the outcomes with current detection techniques to gauge the effectiveness of their methodology. They evaluate several performance indicators, including recall, accuracy, precision, and F1-score, to show how well the deep learning-based approach correctly identifies DDoS attacks.

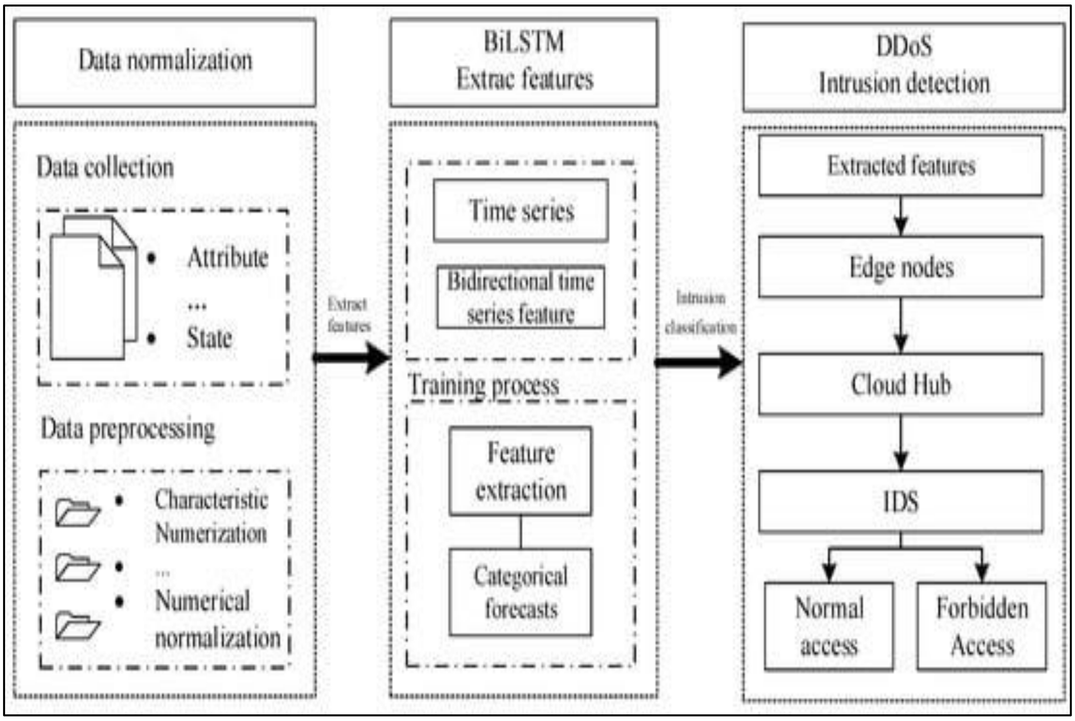


Figure 2.1 frame diagram of the Bi-LSTM model for detecting DDoS [6].

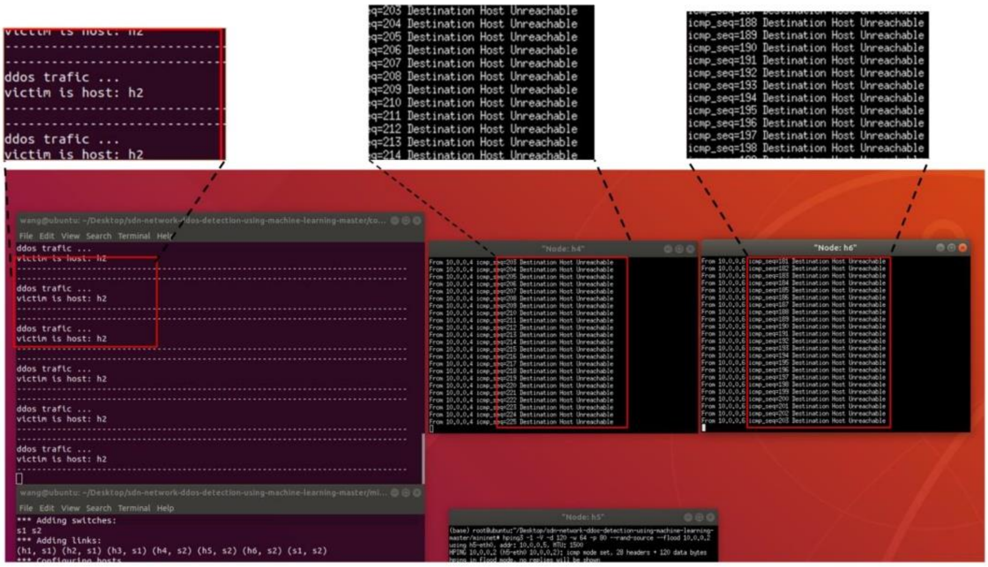
The authors of [12] discuss the increasing danger posed by DDoS attacks and the demand for efficient detection systems. They provide a unique architecture that mixes feedforward and deep neural networks and uses the autoencoder idea to enhance DDoS detection. The suggested method comprises training feedforward and deep neural networks as autoencoders using data from innocuous network traffic. These models learn to identify the underlying patterns and features by reconstructing the input data. Any significant disparities between the reconstructed data and the original input during the detecting phase point to the presence of DDoS assaults. Because SDNs are dynamic and programmable, they are susceptible to various attacks. The authors discuss the difficulty of identifying DDoS attacks in SDNs. To effectively detect DDoS attacks while reducing false positives, they suggest a hybrid technique combining autoencoders' advantages with one-class SVM.

The input data is subsequently rebuilt using the autoencoder, and the reconstruction error is calculated. To differentiate between regular traffic and DDoS attacks, a threshold is established based on the reconstruction error [5]. The recovered data is further classified using a one-class SVM to improve the detection capacity. The SVM learns to recognise variations from typical patterns after training on the normal traffic data. This autoencoder and one-class SVM combo offer a reliable and accurate detection method for DDoS attacks in SDNs. The authors of [13] acknowledge the growing danger posed by DDoS attacks and the demand for efficient detection systems to protect network infrastructures. They investigate the potential of deep learning models and recurrent neural networks (RNNs) for detecting and preventing DDoS attacks. The study begins by giving a general review of DDoS attacks, including their traits and effects on network performance. Additionally, it covers the drawbacks of using conventional detection methods and emphasises the benefits of using RNNs and deep learning models for DDoS assault detection. The authors present long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) models based on neural networks. The LSTM model captures the temporal dependencies and sequential patterns in network traffic data, whereas the CNN model focuses on extracting spatial characteristics from packet payloads.

The study summarises the several deep learning architectures used for DDoS detection. Long Short-Term Memory (LSTM), Autoencoders, Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Generative Adversarial Networks (GANs), and Long Short-Term Memory (LSTM) are all covered. The report describes each method and discusses its advantages for spotting DDoS attacks. The evaluation measures used to gauge how well deep learning models detect DDoS are explained. Accuracy, precision, recall, F1 score, and area under the ROC curve (AUC-ROC) are examples of common metrics that are defined. The study presents results from a few experimental tests that used deep learning techniques to identify DDoS attacks. Results from this research are described, along with the effectiveness of several deep learning models and a comparison of their results [14].

2.3 Overview of DNS and DRDoS Attacks

Software-defined networks (SDNs) are susceptible to serious cybersecurity risks when they are attacked by distributed denial-of-service (DDoS) botnets. In this research,a method is presented for detecting DDoS attacks in software-defined networks (SDNs) using feature engineering and machine learning[2]. First, the CSE-CIC-IDS2018 dataset was cleaned and normalized. Then, an improved binary grey wolf optimization approach was used to find the optimal feature subset. Following this, the ideal feature subset was trained and tested in five different machine learning algorithms: Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbor (k-NN), Decision Tree, and XGBoost. The best classifier was then selected for DDoS attack detection and deployed in the SDN controller. According to the findings, RF achieves the highest levels of performance when compared across a variety of performance criteria (such as accuracy, precision, recall, F1 values, and AUC values). In addition to this, this study investigate how various models and methods compare to one another. The findings indicate that the method that they proposed performed the best and is able to successfully detect and identify DDoS assaults in SDNs. This offers a fresh concept and solution for improving the safety of SDNs.



wang@ubuntu: ~/Desktop/sdn-network-ddos-detection-using-machine-learning-master/co...
File Edit View Search Terminal Help
wang@ubuntu: ~/Desktop/sdn-network-ddos-detection-using-machine-learning-master/mi...
File Edit View Search Terminal Help
"Node: h4"
"Node: h6"
"Node: h5"

```
ddos trafic ...
victim is host: h2
------------------------------------------------------------
------------------------------------------------------------
ddos trafic ...
victim is host: h2
------------------------------------------------------------
------------------------------------------------------------
ddos trafic ...
victim is host: h2
------------------------------------------------------------
------------------------------------------------------------
ddos trafic ...
victim is host: h2
------------------------------------------------------------
------------------------------------------------------------
ddos trafic ...
victim is host: h2
------------------------------------------------------------
------------------------------------------------------------
ddos trafic ...
victim is host: h2
------------------------------------------------------------
```

```
*** Adding switches:
s1 s2
*** Adding links:
(h1, s1) (h2, s1) (h3, s1) (h4, s2) (h5, s2) (h6, s2) (s1, s2)
```

```
From 10.0.0.4 icmp_seq=203 Destination Host Unreachable
From 10.0.0.4 icmp_seq=204 Destination Host Unreachable
From 10.0.0.4 icmp_seq=205 Destination Host Unreachable
From 10.0.0.4 icmp_seq=206 Destination Host Unreachable
From 10.0.0.4 icmp_seq=207 Destination Host Unreachable
From 10.0.0.4 icmp_seq=208 Destination Host Unreachable
From 10.0.0.4 icmp_seq=209 Destination Host Unreachable
From 10.0.0.4 icmp_seq=210 Destination Host Unreachable
From 10.0.0.4 icmp_seq=211 Destination Host Unreachable
From 10.0.0.4 icmp_seq=212 Destination Host Unreachable
From 10.0.0.4 icmp_seq=213 Destination Host Unreachable
From 10.0.0.4 icmp_seq=214 Destination Host Unreachable
From 10.0.0.4 icmp_seq=215 Destination Host Unreachable
From 10.0.0.4 icmp_seq=216 Destination Host Unreachable
From 10.0.0.4 icmp_seq=217 Destination Host Unreachable
From 10.0.0.4 icmp_seq=218 Destination Host Unreachable
From 10.0.0.4 icmp_seq=219 Destination Host Unreachable
From 10.0.0.4 icmp_seq=220 Destination Host Unreachable
From 10.0.0.4 icmp_seq=221 Destination Host Unreachable
From 10.0.0.4 icmp_seq=222 Destination Host Unreachable
From 10.0.0.4 icmp_seq=223 Destination Host Unreachable
From 10.0.0.4 icmp_seq=224 Destination Host Unreachable
From 10.0.0.4 icmp_seq=225 Destination Host Unreachable
```

```
From 10.0.0.6 icmp_seq=181 Destination Host Unreachable
From 10.0.0.6 icmp_seq=182 Destination Host Unreachable
From 10.0.0.6 icmp_seq=183 Destination Host Unreachable
From 10.0.0.6 icmp_seq=184 Destination Host Unreachable
From 10.0.0.6 icmp_seq=185 Destination Host Unreachable
From 10.0.0.6 icmp_seq=186 Destination Host Unreachable
From 10.0.0.6 icmp_seq=187 Destination Host Unreachable
From 10.0.0.6 icmp_seq=188 Destination Host Unreachable
From 10.0.0.6 icmp_seq=189 Destination Host Unreachable
From 10.0.0.6 icmp_seq=190 Destination Host Unreachable
From 10.0.0.6 icmp_seq=191 Destination Host Unreachable
From 10.0.0.6 icmp_seq=192 Destination Host Unreachable
From 10.0.0.6 icmp_seq=193 Destination Host Unreachable
From 10.0.0.6 icmp_seq=194 Destination Host Unreachable
From 10.0.0.6 icmp_seq=195 Destination Host Unreachable
From 10.0.0.6 icmp_seq=196 Destination Host Unreachable
From 10.0.0.6 icmp_seq=197 Destination Host Unreachable
From 10.0.0.6 icmp_seq=198 Destination Host Unreachable
From 10.0.0.6 icmp_seq=199 Destination Host Unreachable
From 10.0.0.6 icmp_seq=200 Destination Host Unreachable
From 10.0.0.6 icmp_seq=201 Destination Host Unreachable
From 10.0.0.6 icmp_seq=202 Destination Host Unreachable
From 10.0.0.6 icmp_seq=203 Destination Host Unreachable
```

```
(base) root@ubuntu:~/Desktop/sdn-network-ddos-detection-using-machine-learning-
master/mininet# hping3 -1 -V -d 120 -w 64 -p 80 --rand-source --flood 10.0.0.2
using h5-eth0, addr: 10.0.0.5, MTU: 1500
HPING 10.0.0.2 (h5-eth0 10.0.0.2): icmp mode set, 28 headers + 120 data bytes
hping in flood mode, no replies will be shown
```

Figure 2.2 Response of the controller to attacks of DdoS [2].

[1] suggested that the number of different kinds of electricity Internet of Things terminal devices has increased by leaps and bounds as a direct result of the rapid growth of smart grids. An assault on one of the end devices that are difficult to safeguard or on any node in a vast and complicated network has the potential to put the grid in danger. Because the traffic that is generated by Distributed Denial of Service (DDoS) attacks is characterized by short bursts of time, it is difficult to apply existing centralized detection methods that rely on human configuration of attack characteristics to changing assault situations. This is because DDoS attacks generate traffic in a manner that is characterized by DDoS attacks. In this study, a DDoS attack detection model that is based on Bidirectional Long Short-Term Memory (BiLSTM) is proposed by constructing an edge detection framework. This structure achieves bi-directional contextual information extraction of the network environment by utilizing the BiLSTM network, and it automatically learns the temporal characteristics of the attack traffic in the original data traffic. The Distributed Denial of Service Attack (DDoS) in Power Internet of Things is the Research Object of This Paper. In terms of accuracy, false detection rate, and time delay, the results of the

simulation reveal that the model surpasses standard advanced models such as Recurrent Neural Network (RNN) and Long Short Term Memory (LSTM). It contributes in a tangential manner to the protection of the power Internet of Things and successfully enhances the dependability of the power grid.

claimed that in the past few decades, there has been an explosion in the use of the Internet, which has led to the installation of high-speed networks in both commercial and educational establishments. The high-speed network is experiencing an increase in the number of security challenges as a direct result of the growing amount of network traffic [3]. Although the Intrusion Detection System (IDS) has a large part in identifying prospective attacks, the enormous traffic flow generates serious technical issues related to monitoring and detecting the activities on the network. These challenges can be attributed to the fact that the IDS has a significant role in recognizing potential assaults. In addition, the destructive character of a Distributed Denial-of-Service (DDoS) assault makes it stand out as a significant cyber-attack independent of the development of Software Defined Network (SDN) architecture due to the fact that DDoS attacks are becoming increasingly common. In this research, a novel framework is proposed to address the performance concerns of IDS as well as the design issues of SDN regarding DDoS attacks.

This is accomplished by putting intelligence in the data layer of the SDN architecture through the utilization of Data Plane Development Kit (DPDK). This innovative architecture is known as the DPDK based DDoS Detection (D3) framework due to the fact that DPDK enables rapid packet processing and monitoring in the data plane. In addition, the DPDK's statistical anomaly detection technique, which is implemented in the data plane as a Virtual Network Function (VNF), enables for the rapid identification of DDoS attacks. The testing results of the D3 framework ensure the unique IDS framework's efficiency as well as its effect. The CIC DoS datasets that are readily available to the public also guarantee that a single statistical anomaly detection technique will have an effective detection effect against a DDoS attack.

In [4] reported that cloud computing provides users with on-demand service options that are delivered through the Internet. The services can be accessed whenever desired and from any location. In spite of the fact that it provides valuable services, the paradigm nevertheless has potential safety flaws. The availability of cloud services can be negatively impacted by a Distributed Denial of Service (DDoS) assault, which also poses a threat to cloud computing's data security. The availability of services for legitimate users is contingent upon the detection of distributed denial of service attacks (DDoS). Numerous academics have put their time and effort into studying this problem, which has resulted in improved accuracy for a variety of datasets. This study will describe a strategy for detecting distributed denial of service attacks in cloud computing. The major goal of this study is to decrease the number of false positives that occur during the DDoS detection process. In the study that is being presented, they choose the features that are most important by employing two different approaches to feature selection. These approaches are known as the Mutual Information (MI) and Random Forest Feature Importance (RFFI) methods. Various machine learning algorithms, including Random Forest (RF), Gradient Boosting (GB), Weighted Voting Ensemble (WVE), K Nearest Neighbor (KNN), and Logistic Regression (LR), are utilized on certain feature sets. The findings of the experiments indicate that RF, GB, WVE, and KNN each have an accuracy of 0.99 when given 19 features to work with. The misclassifications of these methodologies are investigated as part of the ongoing research project, which ultimately results in more precise measurements. The results of numerous testing indicate that the RF performed well in the detection of DDoS attacks, with only a single attack being incorrectly categorized as normal.

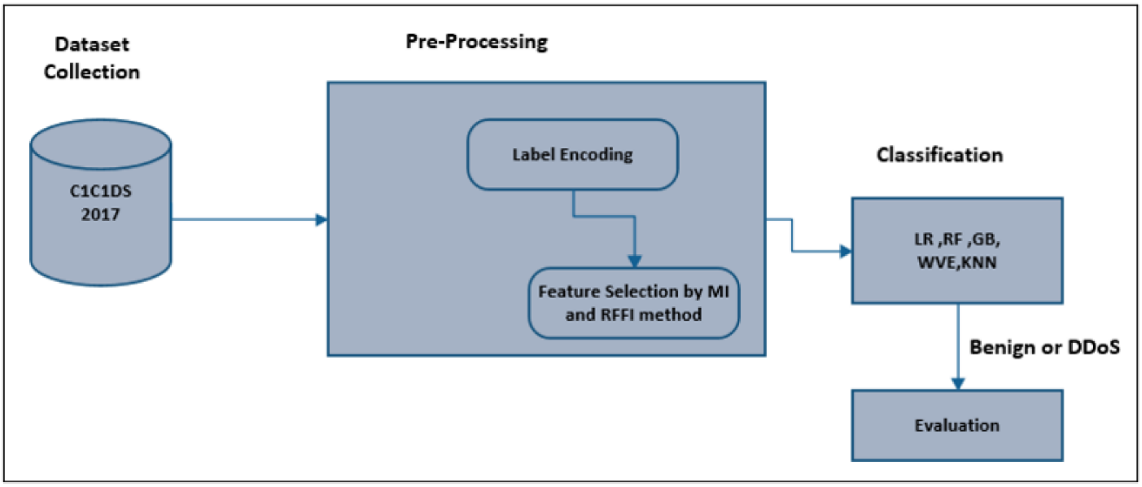


Figure 2.3 Architecture of the proposed DDoS attack detection model [4].

[5] stated that the capacity to recognize and counteract any threat or attack in any network infrastructure, such as a software-defined network (SDN), as well as defend the internet security architecture against various threats or attacks, has significantly increased as a result of recent advancements in security approaches. Among the most widely used methods for thwarting distributed denial-of-service (DDoS) assaults on any sort of network are machine learning (ML) and deep learning (DL). This systematic review's goal is to find, assess, and discuss recent developments in ML/DL-based DDoS attack detection techniques for SDN networks. To accomplish their goal, they carried out a systematic review in which searched for studies that, between 2018 and the beginning of November 2022, identified DDoS attacks in SDN networks using ML/DL methods. they have made considerable use of several digital libraries (including IEEE, ACM, and Springer) and one academic search engine (Google Scholar) to search the contemporary literature. they have examined the pertinent research and divided the SLR's findings into the following five categories: The existing literature covers the following topics: (i) various types of DDoS attack detection in ML/DL approaches; (ii) methodologies, strengths, and weaknesses of existing ML/DL approaches for DDoS attacks detection; (iii) benchmarked datasets and

classes of attacks in datasets; (iv) preprocessing techniques, hyperparameter values, experimental setups, and performance metrics; and (v) current research gaps and encouraging future directions.

2.4 Machine Learning and Deep Learning in Cybersecurity

According to the demand for cybersecurity has grown along with the development of systems that depend more on networking and programming. Cyberattacks pose a changing threat to both organizations and people. Among the destructive cyberattacks, Distributed Denial of Service (DDoS) attacks have become increasingly common among hackers[15]. The hazard is still increasing despite new technologies and safety precautions. An exponential rise in threats has been spurred by the availability of DDoS attack services to anyone with almost no specific set of skills and capabilities.

To lessen the effects, it is crucial to identify these attacks in real time. The CICDDOS2019 dataset is examined and machine learning models to identify DDoS assaults are built as part of this work. To shorten training time, random samples are collected and then feature engineering techniques are performed on top of them. The 15 most important features from a balanced dataset of 360,000 records are extracted. The classification models Decision Tree, Random Forest, Naive Bayes, Stochastic Gradient Boosting, and K Nearest Neighbors are trained and tested on both the original dataset and the balanced dataset. In both datasets, Random Forest produced results with an accuracy of higher than 99%. Evidently, DDoS attacks may be detected in real time with a very high level of accuracy and precision using machine learning techniques.

[16] suggested that in the modern world, technology is more pervasive and accessible than ever across a wide range of platforms and devices, from business servers and commodity PCs to mobile phones and wearables, connecting a wide range of stakeholders, including homes, businesses, and crucial infrastructures.

A huge and complicated danger landscape that is challenging to contain is produced by the sheer quantity and variety of the many operating systems, the device specifics, the varied usage areas, and the accessibility-ready nature of the platforms. It has become increasingly difficult to stay on top of these evolving cyber-threats, which currently heavily rely on gathering and utilizing cyber-threat intelligence before an attack (or at least shortly after, to minimize the damage), and requires the collection, analysis, leveraging, and sharing of enormous volumes of data. In this work,they introduce inTIME, a machine learning-based integrated framework that offers a holistic view of the cyber-threat intelligence process and enables security analysts to quickly identify, gather, analyze, extract, integrate, and share cyber-threat intelligence from a wide range of online sources, including popular social networks, forums, and deep dark web sites. Security analysts and security stakeholders can quickly deploy a variety of data collection services (such as targeted web crawlers, site scrapers, domain downloaders, and social media monitors), automatically rank the collected content according to its likelihood to contain useful intelligence, and identify and extract cyber-threat intelligence and security artifacts via automated to the best of our knowledge, this is the first solution in the literature to offer a complete threat lifecycle via an integrated, user-friendly, yet extendable framework that can support an end-to-end cyber-threat intelligence management platform.

2.4.1 ML/DL Approaches in DRDoS Detection

Due to the assault spectrum that impacts the Data Center servers, high-rate flooding attack detection and categorization has become a crucial component for network administrators. In order to shield web servers from damaging attacks like Distributed Reflection Denial of Service (DRDoS) attacks, the primary goal of this work is to suggest the Protocol Independent Detection and Classification (PIDC) system[17]. The Distributed Denial of Service (DDoS) assault prevention methods are defeated by the DRDoS flooding attack, which takes use of fixed IP spoofing. This is the first study to use SNMP MIB variables to identify and categorize different forms of reflected assaults. Data mining and machine learning techniques are used by the proposed PIDC system to find all kinds of reflected flooding assaults. The Simple Network Management Protocol - Management Information Base (SNMP-MIB) variables used in the rank correlation based detection technique are retrieved from the incoming data, and the algorithm analyzes the relationship between the MIB variables to distinguish between attacks and legitimate traffic. When DDoS flooding attacks are mirrored, the C4.5 classification system extracts and frames association rules based on protocol information. Finally, more resources are distributed to legitimate requests, including more CPU, memory, and disk space. When compared to other reflected attack detection methods, this method obtains 99% true positive rates and a 1% false positive rate. Additionally, these assaults are divided into groups according to their kinds, such as DNS and TCP reflection attacks, which have the highest likelihood of generating attack traffic.

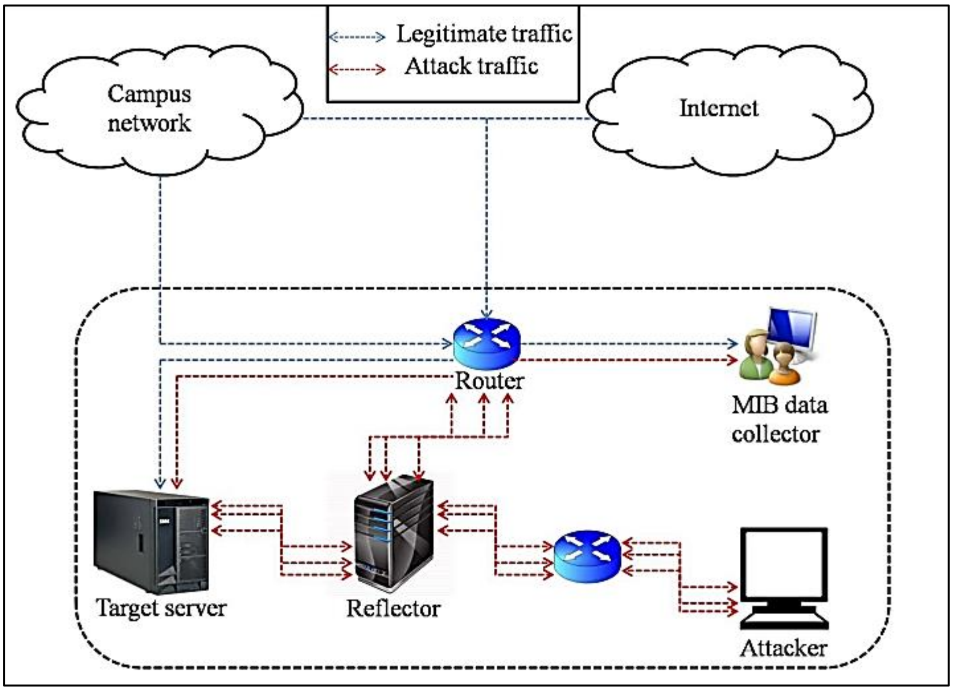


Figure 2.4 Experimental Setup for DRDoS Attack [17].

[18] mentioned that a type of DoS (Denial of Service) attack called a DRDoS (Distributed Reflection Denial of Service) attack connives at getting outside servers to flood the targets with data. That is, in order to transfer data to the victims designated by the source address field of the IP packet, attackers utilize source address IP spoofing to conceal their identity. Reflection is the term used to describe the act of tricking servers of good services into "reflecting" attack traffic to the targets. The majority of currently used detection techniques for these attacks are built around known assaults by protocol, making it challenging to identify unknown ones. their research indicates that there is one protocol-independent detection technique that has been in use. It is based on the premise that the irregular flows from the reflector to the victim have a strong linear relationship. Furthermore, the approach makes the obviously illogical assumption that all packets from reflectors are attack packets when attacked. Five features are discovered to be useful for detecting DRDoS attacks in this study, and they presented a method to detect DRDoS attacks utilizing these features and machine learning methods. Its detection performance is tested experimentally, and the results show that our approach has a considerably higher detection performance.

[9] suggested that the necessity to change the existing network architectures has recently come to the forefront due to improvements in mobile devices and systems, the advent of new concepts like cloud computing and big data, as well as the phenomenal expansion in the number of network users. Software-defined networking (SDN) is one of the methods that shows promise in resolving these issues. Network control and traffic flows are independent of one another and are directly planned in the SDN architecture, which is a singularly new design. Because the SDN's concentrated view of networks is more thorough than previous approaches, it is more effective at fending off hostile attacks, such as amplification attacks. When distributed denial of service (DDoS) attacks are

amplified, the response is more than the request. Amplification attacks send responses to the victim rather than the attacker by using the victim's real address as the source address. Because of this, it is more challenging to find these assaults in conventional networks, whereas SDN's targeted approach can aid in their detection. Machine learning (ML) algorithms are one technique for spotting these threats among many others. In keeping with this, the goal of the current paper is to use ML techniques to detect distributed reflection denial of service (DRDoS) assaults. The simulation is carried out with the aid of ML algorithms, and the results point to a substantial advancement over earlier techniques in the detection of DRDoS attacks.

[19] mentioned that Denial-of-Service (DoS) or Distributed Denial-of-Service (DDoS) attacks, which render online services inaccessible to authorized users by flooding the target system with packets, have become more common in recent years. they suggested two methods in this work to identify Distributed Reflection Denial of Service (DrDoS) attacks in the Internet of Things. To find IoT-DoS attacks, the first way employs a hybrid Intrusion Detection System (IDS). The second method employs deep learning models built on Long Short-Term Memory (LSTM) that have been trained on the most recent dataset for dealing with certain types of DrDoS. Our test findings show that the suggested approaches may identify inappropriate behavior, protecting the IoT network from DoS and DDoS attacks.

2.5 Intrusion Detection System

The Internet is currently subject to a variety of threats that put its data at danger. Therefore, there is a serious concern about the security of the information within the network[8]. The Intrusion Detection System (IDS) was created to detect the outbreak of a stream of attacks and alert the network system administrator providing network security in order to prevent the loss of extremely valuable information. IDS is an extrapolative model used to identify malicious or regular network traffic. The concept of a software-driven network is transformed by the breakthrough paradigm known as Software-Defined Networks (SDN), which separates the control plane from the data plane. SDN gives us the chance to build a managed and programmable network by separating the data and control planes, enabling applications in the top plane to connect to physical devices through the controller. Flow rules are established and network modules are executed by the controller operating in the control plane in order to pass packets to the switches located in the data plane. Cyber attackers target the SDN controller in an effort to take control of the control plane, which is thought of as the brain of the SDN and provides a variety of functionalities such as controlling flow to switches or routers in the data plane below via southbound Application Programming Interfaces (APIs) and business and application logic in the application plane above via northbound APIs to implement complex networks. However, due to its centralized aspect, the control plane becomes a seductive target for security attacks from adversaries. The major studies that are published between 2015 and 2021 that employed Machine Learning (ML) and Deep Learning (DL) approaches to build an IDS solution to provide security for SDN are thoroughly reviewed in this study. In order to ensure the SDN paradigm, they also give two thorough taxonomic studies on IDS, as well as ML-DL methods based on their learning categories. Additionally, they have undertaken some quick study on a few benchmark datasets that were used to build IDS in the SDN paradigm. they present a discussion that clarifies ongoing difficulties and IDS concerns to SDN security to concluding the survey.

[10] argued that during communication in the network environment, the data is vulnerable to numerous threats. Finding network communications intrusions is getting more and more important. In order to create efficient intrusion detection systems, researchers apply machine learning approaches. In

this study, a DDoS intrusion detection system that consists of preprocessing steps and a deep learning model. Convolutional Neural Network, Different Deep Neural Network

and Long Short Term Memory (LSTM)-based models have been tested for this purpose, and their detection and real-time performance have been assessed. The CIC-DDoS2019 dataset, which is extensively used in the literature, is utilized to test the proposed model. The CIC-DDoS2019 dataset underwent preprocessing using methods such feature deletion, random subset selection, feature selection, duplication removal, and normalizing. As a result, the testing and training assessments showed improved recognition performance. According to the test findings, the CNN-based inception-like model produced the best results among the offered models, with accuracy rates of 99.99% for binary and 99.30% for multiclass. Additionally, the proposed model's inference time for various test data sizes appears promising when compared to baseline models with less trainable parameters. When compared to recent studies, the findings from the suggested IDS system and preprocessing techniques are superior.

[11] stated that a new networking paradigm called software-defined networking (SDN) gives the controller centralized control, programmability, and a broad view of the topology. Due to SDN's high audibility, which also poses security and privacy issues, it is growing in popularity. To fend off growing security attacks, SDN needs to be supplied with the greatest security system. A distributed denial-of-service (DDoS) attack uses high-rate packet delivery to flood network channels with erroneous data. Illegal data traffic can overburden network lines, dropping legitimate data and disrupting network services. The Internet, cloud computing platforms, the Internet of Things (IoT), and huge data centers are all vulnerable to the low-rate distributed denial-of-service (LDDoS) attack, a new version of the DDoS attack. Furthermore, because LDDoS attacks

deliver a significant volume of malicious data that is passed off as genuine traffic, they are harder to identify. In order to protect SDN from DDoS assaults, typical security measures like symmetric/asymmetric detection schemes may not be appropriate or ineffective for detecting LDDoS attacks. Therefore, further research investigations in this area are required. There are numerous survey papers that address the DDoS attack detection techniques in SDN, although these research have mostly concentrated on high-rate DDoS attacks. As an alternative, this work presents a thorough analysis of several detection methods proposed to defend the SDN from LDDoS assaults using machine learning techniques. According to our survey, LDDoS attacks can take advantage of vulnerabilities in all tiers of the SDN architecture. Discussions also include current difficulties and potential directions. Researchers can utilize the survey to investigate and create cutting-edge, effective methods to improve SDN's defense against LDDoS attacks.

[20] suggested that in comparison to traditional networks, software-defined networking (SDN) offers programmability, manageability, flexibility, and efficiency. These results from the control and data planes' mutual independence or separation in the SDN. The centralised nature of SDN and the decoupling of two planes improve DDoS attack defense by making it simple to set network device regulations. The controller's global network view is responsible for its capacity to filter network traffic and identify harmful flows. Separating the control and data planes had several advantages, but it also presented a new problem because of its vulnerability to DDoS attacks. One of the most serious risks to SDN is a DDoS assault, in which the culprit interferes with normal users' access to services. In comparison to statistical or policy-based solutions, machine learning (ML) and deep learning (DL) have become effective ways to identify DDoS attacks. A comprehensive taxonomy of DDoS defense techniques has been developed by us. they looked at 260 research studies, and 132 of them

are chosen based on their use of ML and/or DL to identify DDoS attacks in SDN. they cover the previous research that has used feature selection algorithms to choose the best and most effective features for identifying DDoS attacks from a dataset. they outline the characteristics of many publicly accessible DDoS datasets. they also make the case for the necessity of developing datasets specifically for SDN and then employing feature selection techniques that can aid in more effective DDoS attack detection. Finally, they outline the research issues surrounding SDN security that can aid academics in conducting additional studies and creating fresh SDN security strategies.

[21] mentioned that datagrams in the Internet of Things (IoT) are secured by services that provide integrity, secrecy, and authentication. The network is shielded from disruptions and incursions from outside sources. Standard solutions might not work since IoT devices employ a variety of heterogeneous technologies and analyze data over time. Intelligent processes that can be applied to the system's many levels of data flow must be developed. This study uses deep learning-based IDS to look at metainnovations. According to the results of the prior tests, sequential models (LSTM or BiLSTM) are superior for detecting some violent attacks in multiclass classifiers, whereas BiLSTMs are superior for binary (regular/attacker) classification. Deep learning-based intrusion detection systems can now recognize and choose the appropriate structure for each category, according to experts. But in the future, certain issues will need to be resolved. Future efforts should focus on two themes in greater detail. The effect of various data processing methods, such as metamethods or artificial intelligence, on IDS is one of the researchers' main concerns. Among the models, the BiLSTM technique has selected the safest examples with the maximum accuracy. The results show that the BiLSTM architecture is the most trustworthy and practical option for assessing DDoS attacks in IoT.

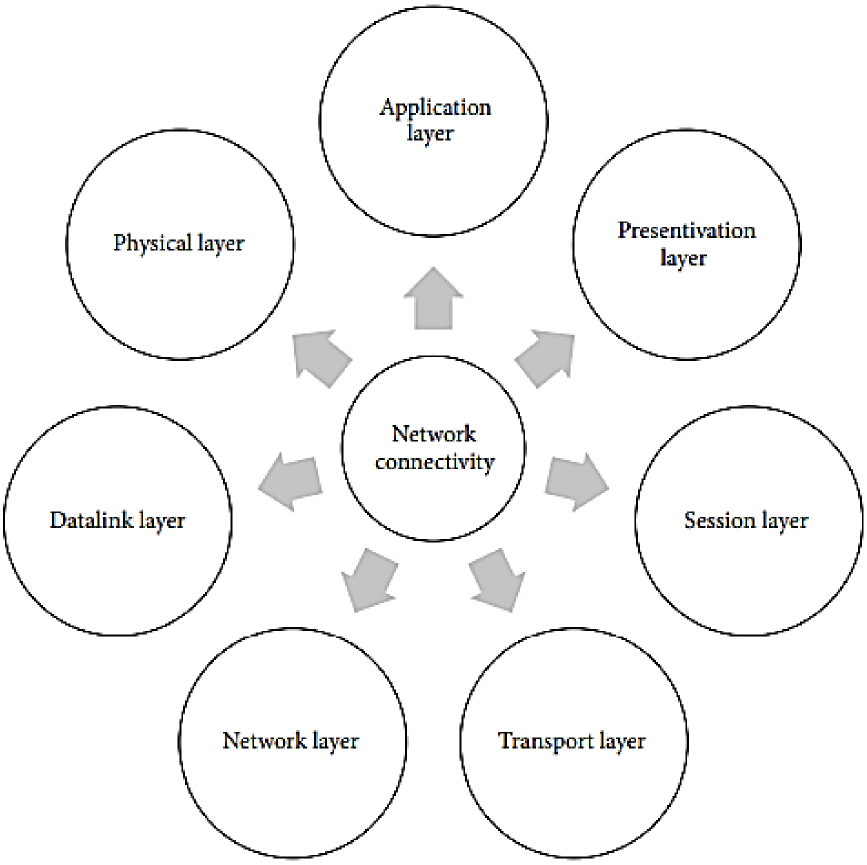


Figure 2.5 The 7-layer conceptual framework for describing network connectivity [21].

[22] stated that it is difficult to identify sophisticated attacks in a number of industries, including business, national defense, and healthcare, as cyberattacks become more intelligent. These sophisticated attacks with unexpected patterns can no longer be detected by conventional intrusion detection systems. Attackers avoid recognized signatures and pose as regular users. An alternative to resolving these problems is deep learning. The list of typical actions or a large number of attack fingerprints are not necessary for Deep Learning (DL)-based intrusion detection to produce detection rules. By using training empirical data, DL creates its own definitions of intrusive features. they create a denial-of-service (DoS) attack-specific DL-based infiltration model. they use the KDD

CUP 1999 dataset (KDD), the most popular dataset for assessing intrusion detection systems (IDS), as the intrusion dataset. KDD includes four different sorts of attacks, including DoS, user-to-root (U2R), remote-to-local (R2L), and probing. Machine learning has been used in numerous KDD studies to divide the dataset into the four categories or into two categories, such as assault and benign. Instead of concentrating on the broad categories, they concentrate on different attacks that fall under the same group. The DoS category of KDD has enough data to train each attack, unlike other KDD categories. they use the most recent IDS dataset, CSE-CIC-IDS2018, in addition to KDD. Compared to KDD, CSE-CIC-IDS2018 uses more sophisticated DoS assaults. In this study, they concentrate on both datasets' DoS categories and create a DL model for DoS detection. they build their model using a convolutional neural network (CNN) and compare its performance to that of an RNN to assess how well it performs. Through a number of studies, they also recommend the best CNN design for improved performance.

[23] mentioned that users would be able to use this article to research the information they need on DDOS attacks globally, forecast upcoming attacks, determine whether their network protection is effective, and assist in debugging it. The goal is to look into potential DDOS assaults, forecast potential attacks on specific IP addresses, the length of the attack, and server load. DDOS attacks worldwide are the focus of the effort. The study on DDOS assaults gathered from all over the world in 2019 is the topic of the work. The primary goal of this effort is to create software that implements the product and machine learning techniques that aid in analyzing and forecasting DDOS attack activity. Based on previous hacker attacks, the algorithm should be able to anticipate attack time, packet volume, server load, and other aspects of DDOS hazards.

[24] reported that the variety and complexity of techniques used to launch distributed denial of service (DDoS) attacks are evolving throughout time. As a result, they provide a way for creating a generalized machine learning (ML)-based model for DDoS attack detection. After examining the different characteristics of the dataset selected for this study, they propose an integrated feature selection (IFS) method that combines two different methods—a filter method and an embedded method—into three stages and selects features that greatly aid in the detection of different DDoS attack types. they train the model for classifying benign and malicious flows using the light gradient boosting machine (LGBM) algorithm. they test the proposed model by sending data of unknown DDoS attack kinds in order to ensure adequate performance and generalized behavior. A variety of performance measures are used to assess the model's performance. they estimate an improvement of about 20% for practically all of the presented measures by comparing the performance of the generated model versus state-of-the-art models. they also demonstrate that a 77% reduction in feature space increases the model's performance. Furthermore, by illustrating a trade-off between high variance and high bias ML models, the created model's generalized behavior is supported.

2.6 Large-Scale Detection using ML and DL approaches

Distributed denial-of-service (DDoS) have become increasingly troublesome in recent years. Therefore, it is imperative to ensure system availability in this ongoing epidemic. For a DDoS multi-classification problem, they present three distinct deep learning algorithms in this study as a network anomaly-based intrusion detection system (N-IDS) [25]. they created a stacked long short-term memory (S-LSTM) neural network, a separate artificial recurrent neural network (RNN), and a deep convolutional neural network (CNN). The third model is a combination of CNN and LSTM. Then, using the most recent flow-based datasets—CICIDS2017, CICDDoS2019, and BoT-IoT benchmarks.

The results show that hybrid CNN-LSTM performs better in practically every validation criterion than the current state-of-the-art techniques.

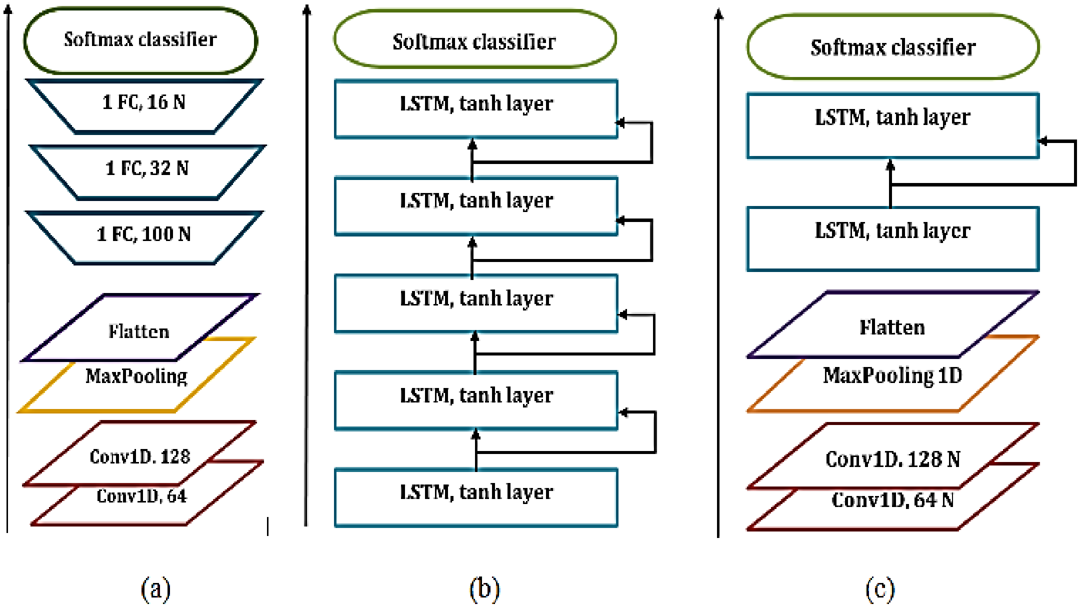


Figure 2.6 The architecture of the three proposed models, (a) deep CNN architecture (b) S-LSTM architecture (c) CNN-LSTM architecture [25].

According to [26], the scholars stated that attacks known as distributed reflective denial of service (DRDoS), particularly those that target open LDAP servers, have grown in popularity in recent years. In these attacks, a brief request for user information is sent to a large number of LDAP servers that are open. As a result, the servers' responses contain substantially more information than what is originally requested, magnifying the traffic and saturating the target with data. Therefore, by applying an upgraded particle swarm optimization (PSO) technique based on an adaptive weighted threshold (AWTPSO) model, this research proposes a unique model for identifying LDAP-based DRDoS assaults. In order to recognize attack patterns, the proposed AWTPSO model integrates aspects of network traffic and LDAP protocol characteristics. The threshold value for each feature is also dynamically adjusted using an adaptive weighted threshold model. The proposed model's detection accuracy is increased by the augmented PSO algorithm's optimization of the threshold values. The recently released CICDDoS2019 dataset (LDAP sub-dataset) has been used to validate the proposed AWTPSO detection model. The experimental results show that, in comparison to other cutting-edge methods, the AWTPSO model detects LDAP-based DRDoS assaults with outstanding accuracy of 99.99% and minimum false positives of 0.01%. As a result, the suggested model offers a very promising and reliable method of identifying the risk of LDAP-based DRDoS attacks on business networks.

Based on study managed by [27], the actual issue with network security is examined in their essay. The problem of recognizing DDoS attacks is being overcome in particular. A solution is put up as part of the study based on broadening the features typically utilized to spot network attacks using a specialized hashing method for certain blocks of device configuration files in the under consideration network of devices. To assure security in the Internet of Things networks, the proposed method is used to identify assaults using machine learning techniques. The CICDDoS2019 dataset is used in the article's comparative comparison of machine learning techniques like Gradient Boosting, AdaBoost, and CatBoost as part of a pilot research. In the instance of binary classification, it is discovered that CatBoost, which has an accuracy of 99.3% and performs on average 0.3% better than the current methods, is the best classifier among those taken into consideration. With an accuracy level of 97%, which is at least 3.9% higher than comparable classifiers, the CatBoost method on a feature set using hashing of data from network devices also exhibits the best performance in the multiclass classification task. The result is unaffected by the multiclass classification's accuracy decline, but it did enable a solution performance improvement of 11.5% when compared to the entire set of features utilized in traditional attack analysis.

[28] reported that networks in the cloud are at serious risk from distributed denial-of-service assaults, or DDoS attacks. Attackers intend to overwhelm the target system with requests and data until it is completely overwhelmed and unable to perform its intended functions. These attacks are constantly improving in sophistication and hazard. One such tactic to evade detection is a low-rate DDoS attack. In the meantime, cloud infrastructure is developing quickly. Cloud computing can utilise resources effectively and scale services in a flexible fashion thanks to container-based technology. When attackers deploy low-rate DDoS attacks, the current approaches for detecting DDoS attacks in cloud computing are insufficient. It is necessary to develop a technique that not only recognizes attacks but also, to some extent, mitigates them. When adversaries employ low-rate DDoS attacks, a Low-Rate DDoS Attack Detection Framework (LRDADF) is put forth for this purpose. Low-rate DDoS attacks are challenging to detect, so a thorough methodology is needed. Along with using deep learning techniques to identify these threats, they also put out a mathematical model to implement a mitigation plan. As a result, they put forth the Hybrid Approach for Low-Rate DDoS Detection (HA-LRDD) as a new algorithm. The algorithm uses a deep auto encoder and convolutional neural networks (CNN) to enable artificial intelligence (AI). Once an attack has been recognized, they devised another technique called Dynamic Low-Rate DDoS Mitigation (DLDM) that lessens its effects. Additionally, it guarantees that the attack is stopped and that the infrastructure keeps running. The suggested framework can detect and mitigate low-rate DDoS assaults to maintain an acceptable level of service in cloud computing settings, according to a thorough simulation study.

The Internet of Things (IoT) is a new age that has been brought about by recent technological innovation according to [29]. Internet-connected items such as smartphones, smart schools, and smart cities are all made possible by modern technology. The advent of smart cities has brought new technological advancements and security issues. One of the most dangerous attacks in the field of network security is the Distributed Denial of Service (DDoS) assault. An efficient BoT-IoT dataset that already exists is employed for this purpose, together with a variety of attack categories and subcategories, for training and evaluating the system's dependability. This study provides an architecture based on the detection of DDoS attack utilizing various machine learning (ML) methodologies in order to improve security in smart cities. Using the BoT-IoT dataset, the paper's main objective is to deploy various machine learning methods, including Random Forest (RF), Naive Bayes (NB), and Decision Tree (DT), to analyze the effectiveness of DDoS attacks. Using the most well-known BoT-IoT dataset, the best accuracy obtained by the machine learning algorithms Random Forest (RF) and Decision Tree (DT) is 91% and 91%, respectively.

A study is led by [30], they mentioned that in that day and age, a person is constantly at risk of a cyberattack. Several technical steps have been developed to stop these cyberattacks or shield people from becoming a target of harmful attacks. Numerous studies have been done on the techniques for detection and prevention. Cyberattack detection heavily relies on machine learning. The distributed denial of service (DDOS) attack is the topic of this study. These malwares operate together to increase erroneous network traffic with the goal of engulfing the targeted website. These types of assaults have been evolving in terms of magnitude, traffic, and modes as technology has developed. Numerous methods, including Random Forest and Convolution Neural Networks, have been used to analyze the study paper's many datasets. In certain studies, the dataset is split into two halves and machine learning methods are used to achieve good precision and accuracy. Two methods are used by certain researchers: one was a mathematical model and the other is a machine learning model. To improve the suggested model's accuracy, resolution time, and precision, a throughput study is conducted using these models.

As a result of an investigation by [31], the authors reported that SDN, or software-defined networking, is emerging as a fresh approach to the growth and innovation of the Internet. SDN is anticipated to be the best option for the Internet's future since it can offer a manageable, flexible, and affordable network. A rare chance to achieve network security in a more effective and flexible way is presented by the introduction of SDN. The centralized controller, the control-data interface, and the control-application interface are SDN's fundamental structural flaws. A variety of assaults can be carried out by intruders using these vulnerabilities. In this study, they provide a deep learning (DL) strategy for an SDN architecture-based network intrusion detection system (DeepIDS). Our models' accuracy for a Fully Connected Deep Neural Network (DNN) and a Gated Recurrent Neural Network (GRU-RNN), when trained and tested on the NSL-KDD dataset, is 80.7% and 90%, respectively. they demonstrate through studies that the DL technique has the capacity to detect flow-based anomalies in the SDN context. they also assess the system's effectiveness in terms of throughput, latency, and resource usage. Our test findings demonstrate that DeepIDS is a workable strategy because it has no negative effects on the OpenFlow controller's performance.

[32] mentioned that as a result, networks are becoming more agile, flexible, and scalable thanks to Software Defined Networking (SDN), the current paradigm in network architecture. Such great characteristics result from the architectural aspect that the control plane in SDN is separated from the data plane and instead resides on a centralized controller with comprehensive knowledge of the network. Security is still a major concern in this area as SDN develops further. The solution to this issue would be considerably aided by an efficient intrusion detection system (IDS), which can monitor real-time traffic, detect, and identify the type of assault. Through the construction of an IDS employing machine learning and genetic algorithm principles, this work seeks to increase

the security of SDN environments. The suggested IDS is split into two stages, the first of which is used to identify assaults and the second of which is used to classify them. These phases are located in the network's switches and controller, respectively. This strategy offers a high attack detection rate while lowering the controller's dependency and workload.

The most frequent and dangerous attacks on both established and cutting-edge networks, including the Internet of Things (IoT), cloud computing, and fifth-generation (5G) communication networks, are distributed denial of service (DDoS) attacks [33]. DDoS assaults have grown increasingly sophisticated and large in recent years. Since it allows for flexibility in both global network monitoring and inline network design, Software-Defined Networking (SDN) technology has proven useful in countering sophisticated threats. A number of papers have suggested ways to identify DDoS attacks, however the majority of them did not make use of the most recent datasets that include the most recent threats. Furthermore, the transition to production networks is made easier by the fact that very few earlier works evaluated their solutions using simulated scenarios. The implementation of a modular and adaptable SDN-based architecture to identify DDoS attacks at the transport and application layers utilizing a variety of machine learning (ML) and deep learning (DL) models is shown in this article. they are able to determine which ML/DL techniques work best under various attack kinds and environmental variables by investigating a variety of techniques. Using two recent security datasets, CICDoS2017 and CICDoS2019, they examined the ML/DL models, and they demonstrated accuracy above 99% when identifying unknown traffic (testing set). With the help of the Open Network Operating System (ONOS) SDN controller and network emulator Mininet, they also set up a test environment. they showed strong detection rates in this experimental system, up to 95% for application-layer DDoS attacks and above 98% for transport-layer DDoS attacks.

2.7 Previous Mitigation Techniques

Modern network approaches like Software-Defined Networking (SDN) separate the control plane from the data plane to provide a flexible design that takes the role of the traditional network architecture [12]. Because of its logically centralized intelligence, programmability, and abstraction, SDN makes network management and monitoring easier. Attacks like a Denial of Service (DoS) attack can be used against SDN architectures. The research on SDN and DoS is reviewed and categorized in this article. The essay also examines the datasets and technologies that are used in the reviewed contributions. In the study, examined methodologies are thoroughly compared in terms of network devices, network layers involved, forms of DoS attacks, and targets of assaults.

[34] suggested that in recent years, there has been a discernible rise in the utilization of cloud computing; consequently, there has been a growth in demand for a variety of cloud platforms like OpenStack, AWS, and others. The provision of trustworthy and risk-free services is one of the difficulties presented by cloud computing. DoS and DDoS assaults, which stand for "denial of service" and "distributed denial of service," have been the most significant challenges to cloud security during the past few years. Hackers launch denial-of-service (DoS) and distributed denial-of-service (DDoS) attacks to overwhelm an online service with an overwhelming volume of traffic coming from a variety of origins. The purpose of this work is to investigate the ways in which the availability of cloud services can be impacted by various types of distributed denial of service (DoS) and distributed distributed denial of service (DDoS) attacks. In addition to that, a number of different preventative measures are presented as well.

TCP SYN Flood is one of the most common kind of denial of service attack that is carried out on computer networks in the modern era according to [35]. they built and distributed modified versions of three network-based mitigation strategies for TCP SYN authentication as a potential countermeasure. This is done in order to prevent further damage. Before forwarding its SYN data, each of them uses the TCP three-way handshake process to create a secure connection with a client. This is done before forwarding its SYN data. These techniques are particularly useful for defending against frequent attacks that use faked IP addresses. Our alterations, on the other hand, make it possible to deflect even more complex SYN floods that are able to sidestep the majority of the conventional techniques. This results in a significantly longer delay for the initial connection attempt; however, subsequent SYN segments see just a negligible increase in latency (less than 0.2 milliseconds). This study presents a comprehensive explanation and analysis of the various techniques, in addition to implementation details and enhancements to security that have been made. CESNET has created a hardware-accelerated FPGA-based DDoS prevention system, which is the basis for the implementations that are being discussed. These implementations are on the verge of being implemented in CESNET's backbone network and the Internet exchange point at NIX.CZ.

According to [36], the scientists mentioned that the smart grid is quickly becoming the industry standard for energy generation and distribution because it integrates cyber-physical systems (CPS) infrastructure with information and communication technologies (ICT). This enables the smart grid to ensure efficient power generation, intelligent energy distribution in real-time, andoptimization. The development of information and communications technology (ICT), on the other hand, has enlarged the attack surface against the energy grid, making it susceptible to a larger variety of cyberattacks. Specifically, denial-of-service, or DoS, attacks could have an effect on the communication network as well as the cyber-physical layer. DoS attacks, which can disrupt the normal operation of genuine smart-grid devices and target a variety of smart grid systems and applications, have emerged as one of the most significant dangers to the smart grid in recent years. In this work, a thorough and methodical overview of DoS assaults in the smart grid is presented. The research examines the most common attack routes and the impact those vectors have on the smart grid. The study also includes a survey of detection and mitigation techniques against DoS assaults in the smart grid using reinforcement learning (RL) algorithms. These techniques include analyzing the strengths and limits of the current approaches and indicating the potential for future research

The impending arrival of the Internet of Things (IoT) has prompted a significant increase in the demand for embedded devices [37]. These devices are designed to enable the independent interaction of sensors and actuators while providing a wide variety of intelligent services. On the other hand, these Internet of Things devices have a limited capacity for compute, storage, and networks, which makes it simple to hack and compromise them. It is vital to design scalable security solutions that are optimized for the ecosystem of the internet of things in order to achieve secure development of the internet of things. To this purpose, software-defined networking (SDN) is a promising paradigm that acts as a pillar in the fifth generation of mobile systems (5G) and has the potential to assist in the detection and mitigation of threats posed by denial of service (DoS) and distributed denial of service (DDoS). In this work, they propose to experimentally evaluate an entropy-based solution to identify and mitigate DoS and DDoS attacks in Internet of Things (IoT) situations utilizing a stateful software-defined networking (SDN) data plane. Specifically, they will be looking at how well this solution works. The findings that are collected show, for the first time, that this method is effective when directed at the data traffic generated by real IoT devices.

The authors in [38] strongly advocate the use of the stochastic back-propagation method for training neural networks, highlighting it as a specific instance of the broader technique known as stochastic gradient descent (SGD). The chapter delves into the background of SGD, elucidating why it's an effective learning algorithm for large training sets, and provides practical recommendations for its implementation.

The authors in [39] address the critical issue of distributed denial of service (DDoS) attacks in cloud computing, a significant threat to the availability of cloud services. They propose a DDoS attack detection system based on an enhanced Self-adaptive evolutionary extreme learning machine (SaE-ELM), which features improved crossover operator adaptability and automatic determination of the number of hidden layer neurons. The system's efficacy is demonstrated through its performance on four datasets, achieving notable detection accuracy and surpassing both the original SaE-ELM based system and other state-of-the-art techniques, albeit with longer training times.

The authors in [40] discuss the ongoing challenge of distributed denial-of-service (DDoS) attacks, particularly in their impact on network security and the exhaustion of target networks with malicious traffic. They propose a hybrid methodology that combines feature selection algorithms with machine learning classifiers for the early detection of DDoS attacks on IoT devices, using the CICDDoS2019 dataset for training and evaluation in a cloud-based setting. The methodology showcases significant performance improvements and feature reduction, highlighting its effectiveness in early DDoS detection.

The authors in [41] explore the increasing sophistication of cyberattacks, with a focus on distributed denial of service (DDoS) attacks. They propose a deep convolutional neural network (CNN) ensemble framework for efficient DDoS attack detection in Software Defined Networks (SDNs), demonstrating improved accuracy over existing detection methods using a current state-of-the-art Flow-based dataset.

The authors in [42] present a new edition of a definitive guide on logistic regression modeling, primarily for health science applications. The expanded Third Edition provides an accessible introduction to logistic regression (LR) models, emphasizing their application in the health sciences. It covers modern statistical software applications, offers rich data sets for practical illustration, and includes new chapters and updated material on various topics, including Bayesian methods and model assessment.

The authors in [43] focus on recent developments in deep neural networks (DNNs), particularly on the optimization of accuracy. They introduce SqueezeNet, a small DNN architecture that achieves AlexNet-level accuracy with significantly fewer parameters and can be compressed to a very small size, making it suitable for deployment on hardware with limited memory. The advantages of smaller DNNs in various applications, such as distributed training and deployment in autonomous vehicles, are also discussed.

The authors in [44] extend the capabilities of semantic instance segmentation by including both visible and occluded parts in their representational output. Their approach involves training a fully convolutional network to produce consistent pixel-level embeddings, enabling accurate estimation of complete masks even in the presence of occlusion, and outperforming traditional bounding-box methods.

The authors in [45] examine the increasing importance of vehicular ad hoc networks (VANETs) in smart transportation systems, emphasizing the critical need for secure and private communication due to the open wireless medium used in VANETs. They provide a comprehensive survey of existing authentication and privacy schemes, comparing them in terms of security, computational and communicational overheads, and resistance to various types of attacks.

The authors in [46] delve into the transformative impact of cloud computing in IT, emphasizing its scalability, virtualization, and cost efficiency. However, they also identify inherent vulnerabilities in the underlying technologies and legacy protocols, particularly susceptible to Distributed Denial of Service (DDoS) attacks. The study introduces a DDoS detection system based on the C.4.5 algorithm, which, in conjunction with signature detection techniques, forms a decision tree for automatic and effective signature attack detection. This system's performance is compared against other machine learning techniques, highlighting its effectiveness in mitigating DDoS threats.

The authors in [47] develop a data-driven model for identifying flow regimes in bubble columns by integrating optical probe technique data with machine learning. They introduce a novel method to determine two key parameters from the optical probe signal—bubble time and characteristic time—providing critical information on operating flow regimes. A machine learning methodology based on support vector analysis is then used to classify flow regimes, demonstrating the ability to accurately categorize different experimental conditions on a single map, which is a significant accomplishment of this research.

The authors in [48] discuss the challenges facing Software-Defined Networks (SDN) in the rapidly evolving landscape of cloud computing, particularly the vulnerability of SDN controllers to Distributed Denial of Service (DDoS) attacks. They propose a deep learning approach utilizing Recurrent Neural Networks (RNNs) for DDoS attack detection on SDN controllers. The approach includes stages of data preprocessing, cross-feature selection, and detection, and is evaluated using benchmark datasets. The results show that this method effectively detects DDoS attacks, with notable accuracy and precision metrics.

The authors in [49] focus on the potential of Software Defined Networking (SDN) to enhance network security and management. They address the limitations of current Machine Learning/Deep Learning intrusion detection systems that rely on supervised learning and well-balanced datasets. To overcome these challenges, they propose a hybrid unsupervised deep learning approach combining stack autoencoder and One-class Support Vector Machine (SAE-1SVM) for Distributed Denial of Service (DDoS) attack detection. Their results demonstrate the algorithm's high accuracy and efficiency, particularly in handling imbalanced and unlabeled datasets.

The authors in [50] highlight the increased reliance on technology during the Covid-19 pandemic and the corresponding rise in Internet-based intrusions and attacks, focusing particularly on Distributed Denial of Service (DDoS) threats. The study conducts a systematic review of deep learning methods for detecting DDoS attacks, analyzing studies from various digital libraries and search engines. The review categorizes findings into several key areas, including types of deep learning approaches for DDoS detection, methodologies and their strengths and weaknesses, benchmark datasets, preprocessing strategies, and future research directions. This comprehensive analysis offers insights into the current landscape of DDoS attack detection and potential advancements.

The authors in [51] discuss the revolution of cloud computing in IT, highlighting its scalability, virtualization, and reduced costs. However, they also point out the vulnerabilities in legacy protocols and underlying technologies that make the system prone to Distributed Denial of Service (DDoS) attacks. To combat this, they propose a novel algorithm called the gradient hybrid leader optimization (GHLBO) algorithm, designed to train a deep stacked autoencoder (DSA) for efficient DDoS attack detection. The method incorporates feature fusion using a deep maxout network (DMN) with an overlap coefficient and data augmentation through oversampling. The GHLBO's performance is evaluated using metrics like the true positive rate (TPR), true negative rate (TNR), and testing accuracy, demonstrating its effectiveness in mitigating DDoS threats.

The authors in [52] address the severity of DDoS attacks in cloud computing environments. They identify limitations in existing IDS for DDoS attack detection, such as delayed convergence and trapping issues. To overcome these, they propose a model combining RNN and deep learning-based strategies, utilizing LSTM and an autoencoder-and-decoder framework with gradient descent learning rule. The model is optimized using a hybrid HHO and PSO algorithm for tuning network parameters and feature selection. The results confirm the superior performance of the proposed LSTM and deep learning model over other models in the literature.

The authors in [53] propose an effective solution for detecting DDoS attacks in cloud servers using a FT-EHO inspired deep belief network (DBN) classifier. The FT-EHO combines Taylor series and elephant herd optimization algorithm with a fuzzy classifier for rules learning. Evaluated using three standard benchmark databases, the FT-EHO's performance is assessed through metrics like accuracy, detection rate, precision, and recall. The results show that the proposed FT-EHO significantly outperforms state-of-the-art methods, demonstrating its efficacy in DDoS attack detection.

The authors in [54] focus on SDN as a key facilitator for agile Internet architecture but also recognize the security issues it presents. They propose a GRU-RNN enabled intrusion detection system for SDNs, tested using the NSL-KDD dataset. Achieving an accuracy of 89% with only six raw features, the results suggest that the GRU-RNN approach does not adversely affect network performance, indicating its potential for effective intrusion detection in SDN environments.

The authors in [55] discuss the ongoing challenges in managing DDoS attacks, emphasizing the difficulty in rapid diagnosis using feature selection algorithms. They propose a hybrid methodology for feature selection, applying methods like chi-square, Extra Tree, and ANOVA on classifiers such as Random Forest, Decision Tree, k-Nearest Neighbors, and XGBoost for early DDoS attack detection on IoT devices. Using the CICDDoS2019 dataset, they demonstrate that the methodology provides superior performance with significant feature reduction and high accuracy, highlighting its effectiveness in early DDoS detection in cloud-based environments.

The authors in [56] explore the use of deep convolutional nets in deep learning, specifically for intrusion detection. They employ a CNN modeling approach, selecting appropriate convolution kernels to extract local correlations and improve feature extraction efficiency. The model, tested on the KDD 99 dataset, shows that it enhances classification accuracy in intrusion detection tasks compared to classical algorithms, indicating the effectiveness of CNNs in this domain.

The authors in [57] highlight the expanding use of SDN and the associated security concerns. They propose a hybrid technique combining deep learning and feedforward neural networks as autoencoders for recognizing DDoS attacks. The model is trained and tested on two datasets, showing high accuracy, precision, recall, and F1-score, both for static and dynamic datasets. The results indicate that the model is highly effective for DDoS attack detection in SDN environments.

The authors in [58] address security issues in cloud-IoT systems, focusing on the challenges posed by centralized web servers in the cloud. They propose a web attack detection system based on distributed deep learning, designed to analyze URLs and deployed on edge devices. The system's performance, tested on several datasets, demonstrates high accuracy and effectiveness in detecting web attacks, suggesting its potential in securing cloud-IoT systems.

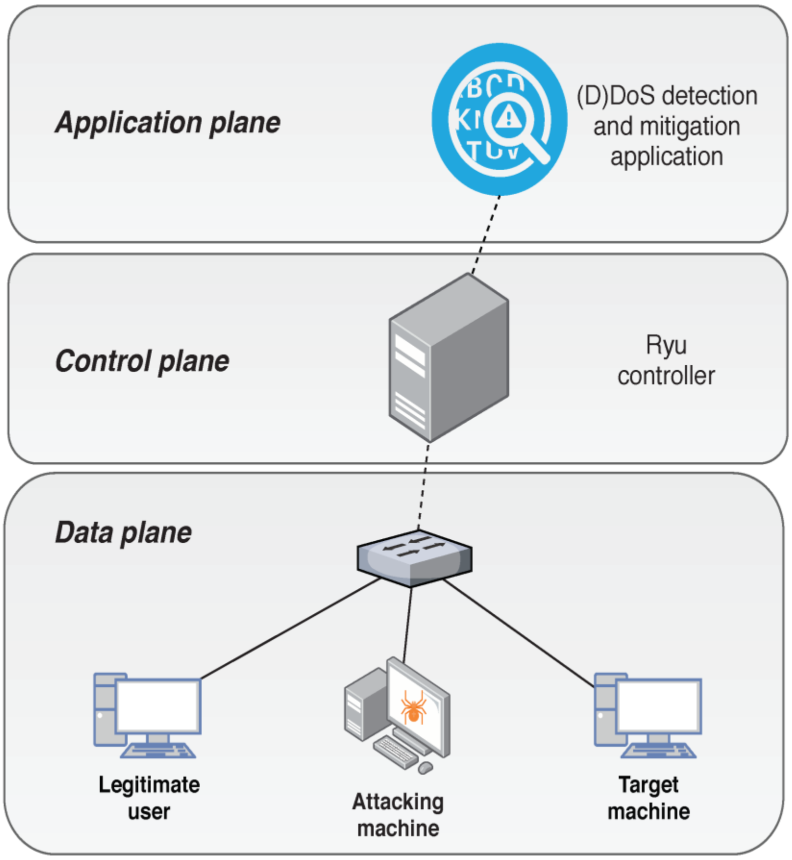


Figure 2.7 Testbed topology of the first scenario [37].

Also, Table 2.1 provides a summary of previous studies mentioned and discussed in this chapter.

Table 2.1 Summary of Previous Studies.

Ref.	Year	Title	Critical Findings	Method
[8]	2021	Intrusion detection system in software-defined networks using machine learning and deep learning, techniques—A comprehensive survey	The critical findings from this survey include: Diverse IDS Techniques in SDNs, Enhanced Detection Accuracy, Adaptability to SDN Dynamics, Feature Engineering	surveyed papers include algorithms like Random Forest, SVM, K-Nearest Neighbors, CNNs, RNNs, and more.
[9]	2022	An Efficient Method for Online Detection of DRDoS Attacks on UDP-Based Services in SDN Using Machine Learning Algorithms.	Critical findings may include the effectiveness and speed of this detection method. Also Critical findings could include insights into which machine learning algorithms are most effective in identifying	Online Learning Algorithms: Online Random Forest, Online Support Vector Machines Other Machine Learning: Random Forest, SVM, Decision Trees Deep Learning Models: Deep neural networks
[10]	2022	A new DDoS attacks intrusion detection model based on deep learning for cybersecurity.	The research findings contribute to the field of cybersecurity by providing a promising approach to bolstering network security against DDoS threats.	This includes specifying the type of neural network (e.g., Convolutional Neural Networks - CNNs, Recurrent Neural Networks - RNNs, or more advanced architectures like Transformer-based models)
[11]	2022	A Survey of Low-Rate DDoS Detection Techniques Based on Machine Learning in Software-Defined Networks	The survey can be used by researchers to explore and develop innovative and efficient techniques to enhance SDN's protection against LDDoS attacks.	algorithms used for this purpose include: SVM, RF, KNN, CNN, RNN, NB and k-means
[12]	2022	A survey on DoS/DDoS mitigation techniques in SDNs: Classification, comparison, solutions, testing tools and datasets.	Offers a holistic view of DoS/DDoS mitigation in SDNs and suggests future research directions in this vital field of cybersecurity.	the paper explores innovative solutions tailored for SDN environments and addresses the availability of testing tools and datasets for evaluating these techniques.
[5]	2023	Machine Learning Techniques to Detect a DDoS Attack in SDN: A Systematic Review.	the main findings and key insights from the selected studies. Analyze trends in machine learning techniques, algorithm performance, and their effectiveness in detecting DDoS attacks in SDN	various machine learning algorithms, such as Support Vector Machines (SVM), Random Forest, and Neural Networks

[20]	2023	A Comprehensive Analysis of Machine Learning-and Deep Learning-Based Solutions for DDoS Attack Detection in SDN.	The findings of this comprehensive analysis reveal insights into the strengths and weaknesses of various machine learning and deep learning-based DDoS detection approaches.	Support Vector Machines (SVM) Random Forest K-Nearest Neighbors (K-NN) Deep Learning (e.g., Convolutional Neural Networks - CNNs, Recurrent Neural Networks - RNNs) Naive Bayes Clustering Algorithms (e.g., K-Means, DBSCAN)
[34]	2016	A study on the impacts of DoS and DDoS attacks on cloud and mitigation techniques.	The study's results would include findings related to the impacts of DoS and DDoS attacks on cloud environments and the effectiveness of the mitigation techniques.	Traffic Analysis Algorithms ,Machine Learning Algorithms ,Rate Limiting Algorithms, Filtering Algorithms ,Intrusion Detection Algorithms
[15]	2022	Machine learning algorithms for DDoS attack detection in cybersecurity.	Maximum accuracy was achieved using Random Forest in both datasets, with an accuracy of more than 99%.	Random Forest Support Vector Machines (SVM) K-Nearest Neighbors (K-NN) Decision Trees Naive Bayes Deep Learning Models (e.g., Convolutional Neural Networks – CNNs)
[21]	2022	Ml-ddosnet: Iot intrusion detection based on denial-of-service attacks using machine learning methods and nsl-kdd.	the testing samples' accuracy for 30% of the data is 79.5 percent. Furthermore, the sensitivity, specificity, and accuracy scores are 97.9 percent, 67.3 percent, and 66.5 percent, respectively.	Random Forest Support Vector Machines (SVM) K-Nearest Neighbors (K-NN) Decision Trees Naive Bayes Deep Learning Models (e.g., Convolutional Neural Networks - CNNs, Recurrent Neural Networks - RNNs)
[37]	2020	Detection and mitigation of DoS and DDoS attacks in IoT-based stateful SDN: An experimental approach.	In this work, the study propose to experimentally evaluate an entropy-based solution to detect and mitigate DoS and DDoS attacks in IoT scenarios using a stateful SDN data plane. The obtained results demonstrate for the first time the effectiveness of this technique targeting real IoT data traffic.	Machine Learning Algorithms ,SDN Controller Logic, Flow-Based Analysis
[18]	2016	A machine learning based approach for detecting DRDoS attacks and its performance evaluation.	In this study, found five features are effective for detecting DRDoS attacks, and this work proposed a method to detect DRDoS attacks using these features and machine learning algorithms	Random Forest Support Vector Machines (SVM) K-Nearest Neighbors (K-NN) Decision Trees Naive Bayes Ensemble Methods Deep Learning Models (e.g., Convolutional Neural Networks - CNNs, Recurrent Neural Networks - RNNs)

[35]	2021	Defense against syn flood dos attacksˇ using network-based mitigation techniques.	The discussed implementations are built on top of the hardware-accelerated FPGA-based DDoS protection solution developed by CESNET and are about to be deployed in its backbone network and Internet exchange point at NIX.CZ.	Traffic Analysis Algorithms Mitigation Algorithms
[22]	2020	CNN-based network intrusion detection against denial-of-service attacks.	The experimental results of binary classification for 18 experimental scenarios show that most of scenarios have more than 99% of accuracy.	Convolutional Neural Network (CNN) Architecture
[16]	2021	intime: A machine learning-based framework for gathering and leveraging web data to cyber-threat intelligence.	identify and extract cyber-threat intelligence and security artifacts via automated natural language understanding processes	• Natural Language Processing (NLP) techniques for text analysis • Classification algorithms for threat categorization • Clustering algorithms for identifying patterns in threat data • Time-series analysis for threat trend prediction
[2]	2023	A DDoS Detection Method Based on Feature Engineering and Machine Learning in Software-Defined Networks.	the RF classifier performed best under accuracy, precision, recall, and f1_score metrics with 0.9913, 0.9843, 0.9992, and 0.9913, respectively.	• Random Forest • Support Vector Machines (SVM) • K-Nearest Neighbors (K-NN) • Decision Trees • Naive Bayes • Deep Learning Models (e.g., Convolutional Neural Networks - CNNs, Recurrent Neural Networks - RNNs)
[23]	2020	DDOS Attacks Analysis Based on Machine Learning in Challenges of Global Changes.	The main task of this work is to develop software implementation of the product, machine learning methods that will help to investigate and predict the activities of DDOS attacks.	• Random Forest • Support Vector Machines (SVM) • K-Nearest Neighbors (K-NN) • Decision Trees

2.8 Chapter Summary

thorough reviewing and conducting that reveals over the past decade, during the digital age, global online platforms, communication networks, and social media sites have experienced a significant rise in cyberbullying incidents. These incidents encompass various forms such as hate speech, offensive language, and inappropriate content on platforms like Facebook, Twitter, and other web applications that involve harassment. As a result, numerous peer-reviewed articles and academic publications have extensively discussed the valuable contributions and benefits of employing ML models and intelligent AI techniques to actively address and mitigate the detrimental consequences of cyberbullying on Twitter. Furthermore, the reviewed studies highlight the significant roles of sentiment analysis and NLP in effectively handling such issues. These practical techniques have demonstrated their efficacy in detecting and preventing harmful online threats.

Building on this discussion, this work is conducted to provide more numerical analysis on the critical role of modern ML models and intelligent algorithms in detecting cyberbullying problems in the Twitter platform, focusing on new relevances and vital contributions of this aspect when AED is integrated into the detection process.

Chapter three is going to discuss and highlight the major research phases and numerical analysis steps followed and adopted to analyze the cyberbullying problem with the consideration of the AED, to find out whether the adoption of this concept could enhance the reliability, performance, efficacy, and accuracy of cyberbullying detection process compared with lower accuracy ML models.

Chapter Three

Methodology And Approaches

3.1Introduction

This chapter, introduces a meticulously designed framework that implements an end-to-end methodology for the detection of DDoS attacks, anchored on a robust dataset specifically curated for DrDoS DNS assault analysis. This dataset is a treasure trove of critical attributes, each playing a pivotal role in the deep understanding and pinpointing of DDoS attack patterns, forming the core from which our model draws its strength. Upon acquiring the dataset, our first course of action is to undertake an extensive Exploratory Data Analysis (EDA). This crucial step allows us to delve into the dataset's intricacies, spotlight key features for attack identification, and eliminate any data irregularities or outliers that could skew our results. Armed with the knowledge gained through EDA, transition into the data preprocessing phase. Here, by applying a suite of sophisticated techniques, including data scaling and normalization, to refine our dataset. This careful preparation is essential, ensuring that the data is in the ideal state for the subsequent Machine Learning (ML) processes, setting the stage for a robust and reliable DDoS detection system.

3.2Proposed Approach

Our proposed framework adopts an quit-to-quit method for DDoS attack detection, drawing its foundational statistics from the specialized DrDoS DNS attack dataset. This dataset, replete with a mess of attributes critical for information and figuring out DDoS conduct, serves as the primary fuel for our version. once the information is acquired, interact in thorough Exploratory records evaluation (EDA) to analyze the dataset's characteristics, highlight important capabilities, and root out any inconsistencies or anomalies. knowledgeable by way of insights gained from EDA, one proceeds to preprocess the statistics, using strategies along with scaling and normalization, to make certain it's far optimized for ML tasks. After these preliminary stages, the facts is then partitioned into education and test subsets. This bifurcation lets in us to appoint a number of ML algorithms along with LR, SVM, and SGD in addition to DL strategies like CNN, AlexNet. via this multi-algorithmic, layered approach, our version pursuits to serve as a comprehensive, correct, and scalable solution for detecting DDoS assaults correctly.

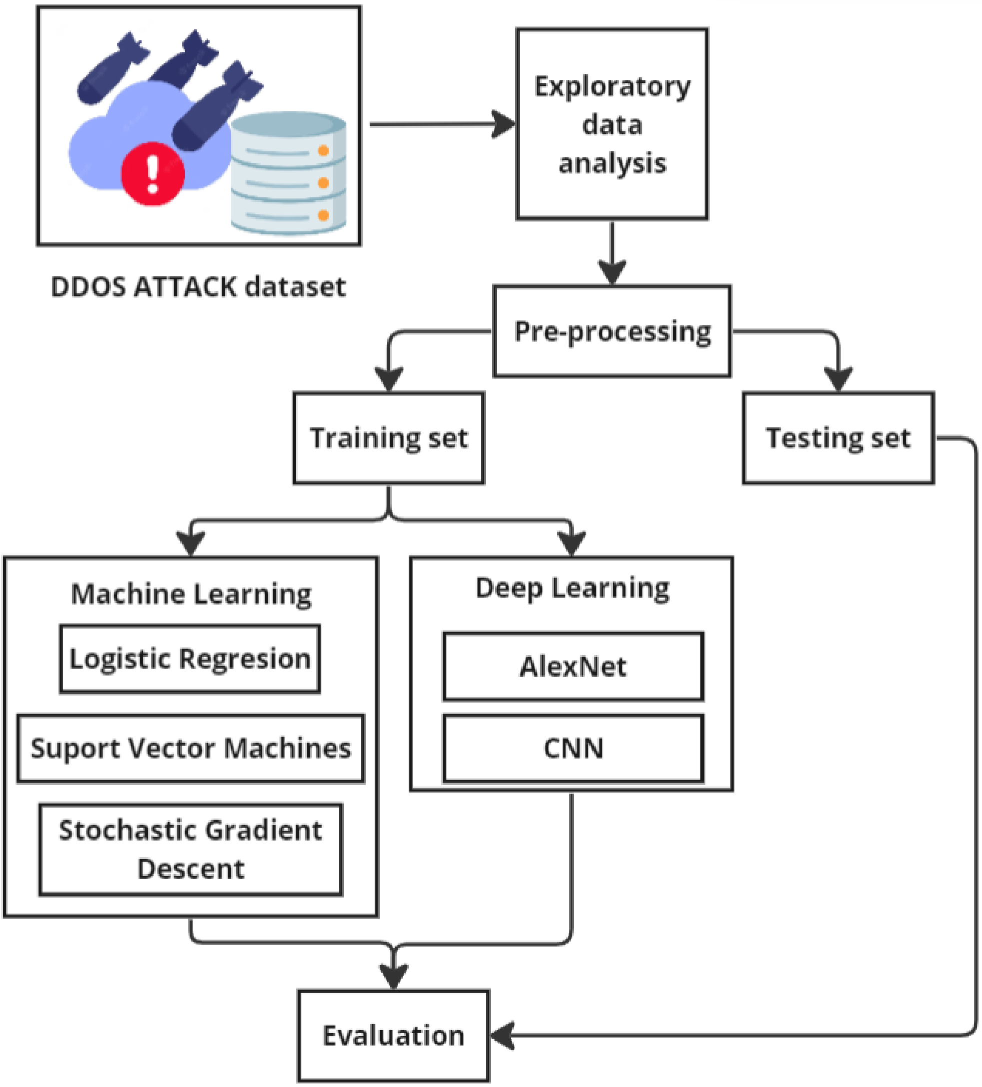


Figure 3.1 Frame diagram of DDOS classification approach

3.2.1 Dataset Description

The "DrDoS DNS Attack" dataset available on Kaggle1 serves as an invaluable resource for researching Distributed Reflective Denial of Service (DrDoS) attacks, particularly focusing on DNS-based exploits. The dataset consists of 33,925 rows and 16 columns, each capturing various metrics and features essential for understanding network behavior and identifying malicious activity. The columns in the dataset cover a range of variables including:

- **Protocol:** Indicates the type of protocol used in the network flow.

- **Flow Duration:** Records the total time duration of the flow in milliseconds, providing an overview of how long the connection lasted.

- **Total Forward Packets:** Counts the total number of packets sent from the source to the destination.

- **Total Backward Packets:** Counts the total number of packets sent from the destination back to the source.

- **Total Forward Packets Length:** Measures the total length of packets sent from the source to the destination in bytes.

- **Total Backward Packets Length:** Measures the total length of packets sent from the destination back to the source in bytes.

[1] https://www.kaggle.com/datasets/lotfikamel/drdos-dns-attack

- **Forward Packet Length Mean:** Provides the average size of the forward packets, offering an idea of the typical packet size going from source to destination.

- **Backward Packet Length Mean:** Provides the average size of the backward packets, offering an idea of the typical packet size going from destination to source.

- **Forward Packets Per Second:** Measures the rate at which packets are sent from the source to the destination, giving insights into the intensity of the flow.

- **Backward Packets Per Second:** Measures the rate at which packets are sent from the destination back to the source, providing insights into the responsiveness of the destination.

- **Forward IAT Mean:** Represents the mean inter-arrival time of forward packets, shedding light on the flow's regularity from source to destination.

- **Backward IAT Mean:** Represents the mean inter-arrival time of backward packets, shedding light on the flow's regularity from destination to source.

- **Flow IAT Mean:** Averages the inter-arrival times for all packets in the flow, providing an overall view of the flow regularity.

- **Flow Packets Per Seconds:** Measures the overall rate of packets in the flow, offering a holistic view of flow activity.

- **Flow Bytes Per Seconds:** Measures the overall rate of data transfer in the flow, complementing 'Flow Packets Per Seconds' for a more comprehensive view.
- **Label:** Indicates whether the flow is malicious or benign.

3.2.2Exploratory data analysis (EDA)

In the "DrDoS DNS Attack" dataset, the correlation matrix provides valuable insights into how different features relate as shown in figure 3.2. One interesting observation is that the 'Protocol' doesn't seem to correlate numerically with other features, likely because it's a categorical variable. Features like 'flow duration' have a pronounced positive correlation with 'forward iat mean' (0.7536) and 'flow iat mean' (0.7401), hinting that longer flow durations often coincide with longer packet inter-arrival times.

Furthermore, the 'total forward packets' feature aligns almost perfectly with 'total forward packets length' (0.999). This means that as packet numbers grow, their combined length does too, almost in tandem. In contrast, benign flows, as opposed to malicious ones, tend to have more'total backward packets', as reflected by its strong negative correlation with 'label' (-0.993). There's also a clear connection between 'backward packet length mean' and 'total backward packets length' (0.998), showcasing that the size of individual packets plays a big role in determining the total length of backward packets. Another observation is that high rates of backward packet transmissions frequently come with more packets, as seen from the high correlation between 'backward packets per second' and 'total backward packets' (0.9549). On the flip side, there are some negative correlations to consider. For example, 'forward packet length mean' tends to decrease as 'total forward packets' increase, given their correlation of -0.326.

Also, benign flows often have more backward packets and a faster transmission rate, as indicated by the strong negative correlations between 'label' with 'total backward packets' (-0.993) and 'backward packets per second' (-0.953). Another captivating detail is the tight positive correlation (0.853) between features like 'flow packets per seconds' and 'flow bytes per seconds'. This implies that when there's a surge in the rate of packet transmission, those packets usually contain more data.

Lastly, the 'label' feature, which is our main focus, has noteworthy positive correlations with 'total forward packets' (0.1315) and 'total forward packets length' (0.1397). But it contrasts sharply with most 'backward' metrics, especially 'total backward packets' (-0.993) . This suggests that these backward-related features play a crucial role in determining whether a flow is malicious or benign.

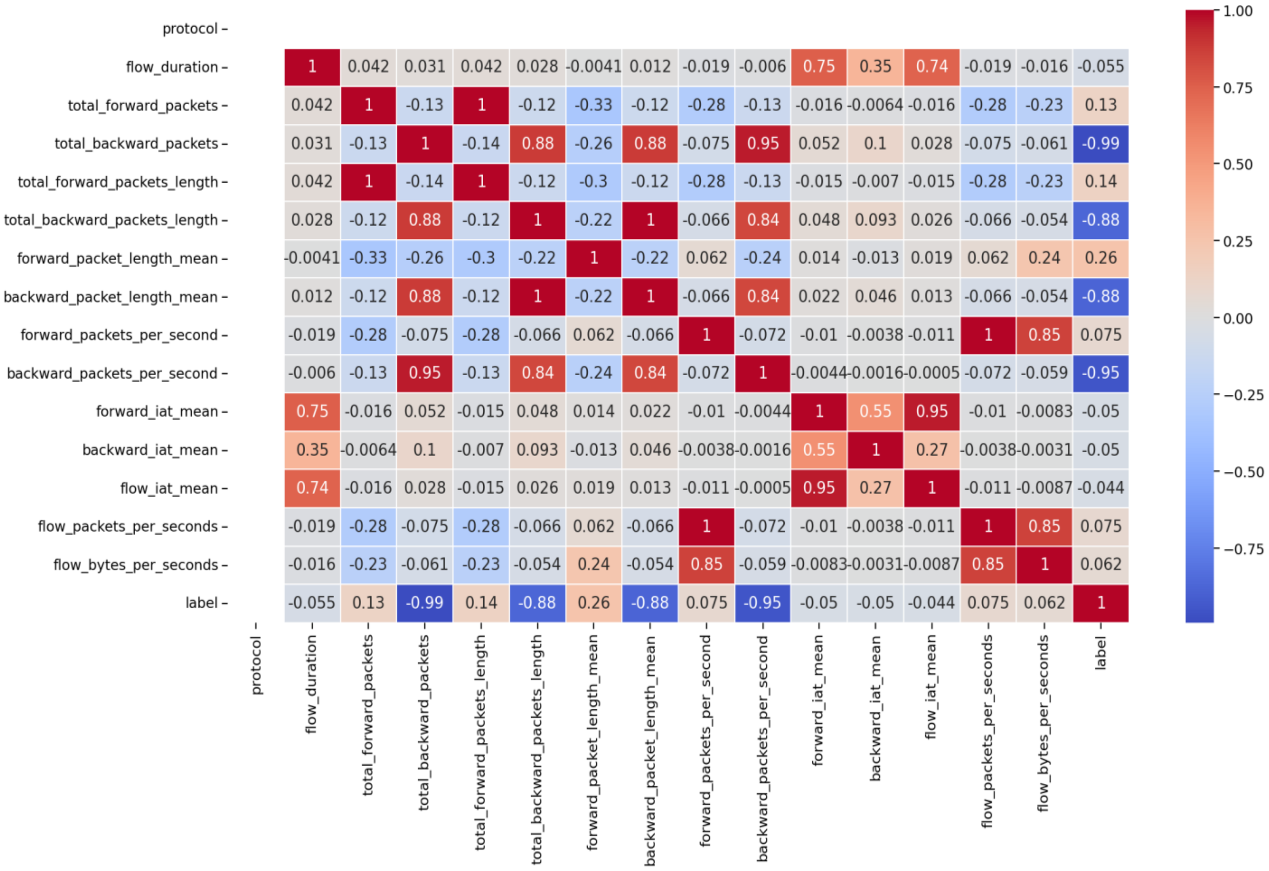


Figure 3.2. Data Correlation

Delving into the "DrDoS DNS Attack" dataset, its descriptive statistics serve as a window into the patterns and traits scattered across various features. One striking fact is the uniformity of the "protocol" feature: all its statistical metrics, from the minimum to the maximum, hover around 17. This implies that there's no diversity in protocol types within the dataset, rendering it more or less redundant for machine learning endeavors as shown in table 3.1.

The "flow duration" displays a substantial spread, with its standard deviation (1.83e+06) dwarfing its mean (8.60e+04). Notably, the max value leaps well beyond the 75th percentile, hinting at potential outliers or a skewed dataset.

Similarly, features like "total forward packets" and "total forward packets length" possess means that overshadow their medians, suggesting a rightward tilt in their distributions. This trend is mirrored,in"total backward packets"and "total backward packets length," albeit with noticeably lower mean values. Interestingly, both "forward packet length mean" and "backward packet length mean"

average values that are distinct from their zero medians unveil distributions that might be heavily skewed, peppered with numerous zeroes. rate-oriented features, such as "forward packets per second" and "backward packets per second," exhibt hefty standard deviations spotlight the diversity inherent in the dataset.

A pronounced gap between the max value and the 75th percentile in these metrics also hints at the lurking outliers. Shifting our gaze to features that capture inter-arrival times - "forward iat mean," "backward iat mean," and "flow iat mean" significant standard deviations and a wide range of values from min to max are observed, indicating a diverse data landscape. Considering "flow bytes per seconds" and "flow packets per seconds," their substantial standard deviations reinforce earlier observations. the dataset exhibts considerable variability, potentially including outliers.

Table 3-1 Data Analysis

	Count	mean	std	min	25%	50%	75%	max
protocol	33925.0	1.700000e+01	0.000000e+00	17.000000	1.700000e+01	1.700000e+01	1.700000e+01	1.700000e+01
flow_duration	33925.0	8.597836e+04	1.831408e+06	1.000000	4.400000e+01	2.350000e+02	2.911900e+04	1.183569e+08
total_forward_packets	33925.0	6.551328e+01	8.938778e+01	2.000000	2.000000e+00	2.000000e+00	1.780000e+02	4.000000e+02
total_backward_packets	33925.0	6.596905e-02	3.581863e-01	0.000000	0.000000e+00	0.000000e+00	0.000000e+00	4.000000e+00
total_forward_packets_length	33925.0	2.940005e+04	3.887001e+04	0.000000	9.600000e+02	2.672000e+03	7.788800e+04	1.760000e+05
total_backward_packets_length	33925.0	6.161120e+00	3.792449e+01	0.000000	0.000000e+00	0.000000e+00	0.000000e+00	7.560000e+02
forward_packet_length_mean	33925.0	6.422248e+02	4.348924e+02	0.000000	4.361600e+02	4.400000e+02	9.320000e+02	1.472000e+03
backward_packet_length_mean	33925.0	3.072071e+00	1.889888e+01	0.000000	0.000000e+00	0.000000e+00	0.000000e+00	3.780000e+02
forward_packets_per_second	33925.0	1.792975e+05	4.423623e+05	0.039297	5.604045e+03	9.478673e+03	4.545455e+04	4.000000e+06
backward_packets_per_second	33925.0	2.745483e+00	1.552335e+01	0.000000	0.000000e+00	0.000000e+00	0.000000e+00	9.815469e+01
forward_iat_mean	33925.0	9.789976e+03	3.937221e+05	0.000000	3.800000e+01	9.936683e+01	2.270000e+02	3.392962e+07
backward_iat_mean	33925.0	1.991698e+03	2.151055e+05	0.000000	0.000000e+00	0.000000e+00	0.000000e+00	2.807712e+07
flow_iat_mean	33925.0	8.961677e+03	3.424422e+05	0.333333	4.400000e+01	1.276667e+02	2.390000e+02	3.392962e+07
flow_packets_per_seconds	33925.0	1.793002e+05	4.423612e+05	0.039297	5.604045e+03	9.478673e+03	4.545455e+04	4.000000e+06
flow_bytes_per_seconds	33925.0	1.270537e+08	3.820510e+08	0.000000	2.358741e+06	5.449393e+06	4.086275e+07	2.944000e+09

3.2.3Data Pre-processing

An essential aspect of any ML pipeline is data preprocessing, often cited as a critical step in determining the success of the model. our focus is on tailoring preprocessing steps to cater to the specificities of our DrDoS DNS Attack dataset. A standout quality of our dataset is its complete absence of missing values. This is a real boon since it spares us from resorting to imputation methods that might inadvertently muddy the waters with bias or added noise. Therefore, our preprocessing is laser-focused on priming the features for optimal machine learning outcomes.

The dataset comprises a mix of numerical attributes that are recorded in different scales and units. ML and DL algorithms, are sensitive to the scale of input features, making feature scaling indispensable. For this reason, the Standard Scaler from the scikit-learn library is employed to normalize the feature set. The Standard Scaler standardizes the features by removing the mean and scaling to unit variance. In mathematical terms, for each feature, the mean is subtracted from each data point and the result is divided by the standard deviation. The process can be represented as:

$$X' = (X-\mu) / \sigma \qquad (3\text{-}1)$$

Where, X is the scaled feature, X is the original feature, μ is the mean, and σ is the standard deviation. After scaling, each feature in the dataset will have a mean of 0 and a standard deviation of 1, thereby ensuring numerical stability and improving the algorithmic performance. The features (X) and labels (y) are then separated to form the input and output for our ML models.

3.2.3.1 Data Splitting

Following the preprocessing stage, one of the pivotal aspects of our pipeline involves segregating the data into distinct training and testing sets. In accordance with best practices, by designating 80% of the data for the training phase, reserving the remaining 20% explicitly for testing. This bifurcation fulfills two essential criteria: On the one hand, the generous allocation to the training set ensures that our machine and DL algorithms have an ample data pool to learn from, facilitating a nuanced understanding of the underlying features and patterns essential for DDoS detection.

On the other hand, setting aside a 20% subset for testing enables a rigorous evaluation of the model's efficacy on data it has not previously encountered. This 80-20 split is intentionally designed to achieve a harmonious balance between comprehensive training and unbiased assessment, ultimately enhancing the model's capacity for dependable and broadly applicable DDoS attack detection.

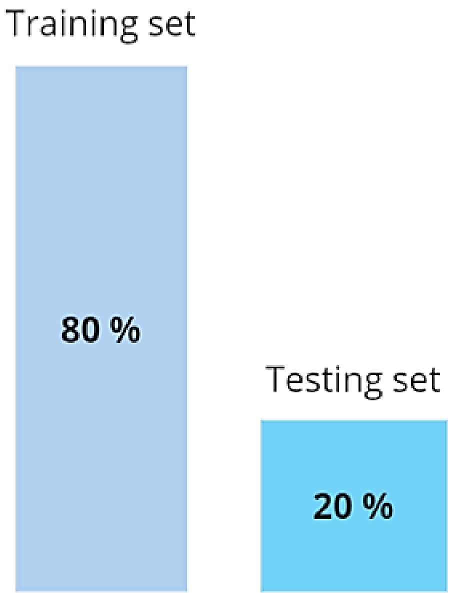


Figure 3.3 Data Splitting

3.2.4Machine Learning models

In the proposed model for this thesis, various ML algorithms are leverarged to address the complexities of our target problem. Each algorithm ranging from LR and SVM to SGD, CNN, and AlexNet has been meticulously selected and tuned to contribute a unique strength to the overall architecture.

3.2.4.1 Logistic Regression method

In the context of this thesis, LR serves as a ML model specifically designed for binary classification tasks, enabling the prediction of a categorical dependent variable based on one or more independent variables

[9] . Unlike linear regression, which outputs continuous values, LR uses the logistic function to squeeze the output probabilities between 0 and 1. The architecture of the LR model comprises several components that contribute to its flexibility and utility.

The core of the model is the logistic function, often referred to as the sigmoid function, defined as:

$$\boldsymbol{f(x)} = \frac{\mathbf{1}}{\mathbf{1+e\ -x}} \qquad \textbf{(3-2)}$$

This function takes a linear combination of the input features X and the weight vector W, along with an intercept term b, and maps this combination into a probability score between 0 and 1. Mathematically, this is expressed as:

$$\boldsymbol{P(Y=1)} = \frac{\mathbf{1}}{\mathbf{1+e}^{\mathbf{-(W\cdot X+b)}}} \qquad \textbf{(3-3)}$$

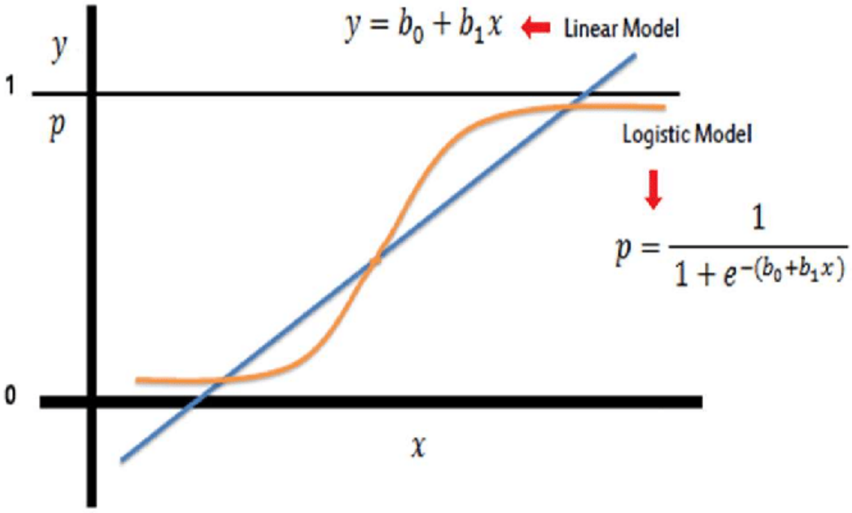


Figure 3.4 Logistic Regression [6].

Regarding hyperparameters, the regularization strength C plays a critical role in controlling the complexity of the model. A smaller value of C increases the regularization effect, helping to mitigate overfitting by discouraging overly complex models. In this study, a C value of 0.03 is selected after a series of experiments aimed at finding an optimal trade-off between bias and variance. Solver algorithms are an often-underestimated component in the architecture of LR models. on this study, delving into algorithms that are tasked with refining the price feature to pinpoint the first-class parameters (W,b) for our model. meticulous examination is conducted on 5 awesome solvers 'newton-cg', 'lbfgs', 'liblinear', 'sag', and 'saga' to gauge their impact at the model's precision. each solver employs its precise mathematical approach to reach at the appropriate answer. the choice of solver could make a full-size distinction, potentially main to more correct results and speedier training levels. The model is trained using a particular dataset and then tested its adaptability on a one of a kind dataset. various performance indicators, like accuracy, for each solver. This allowed us to determine which algorithm stood out the maximum for the dataset to hand.

3.2.4.2 Support Vector Machine (SVM) method

In this thesis, SVM are rigorously investigated as a powerful tool for category tasks [10]. SVM works by way of locating the hyperplane that best divides a dataset into lessons. The core precept behind SVM is to maximize the margin among different instructions within the function space, efficiently enhancing the model's generalization overall performance. One of the version's most first rate attributes is its flexibility, which is done through the use of different kernel features that permit it to perform properly on both linear and non-linear problems [59].

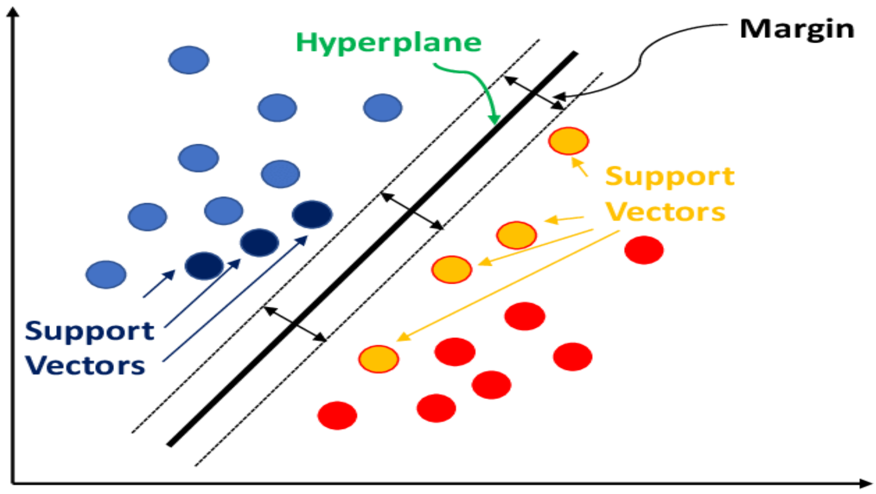


Figure 3.5 . Support Vector Machines [7].

In terms of architecture, the kernel function serves as the cornerstone of the SVM model. A kernel function implicitly computes the dot product among two observations in a higher-dimensional area with out actually reworking the data, thereby permitting the model to discover a hyperplane on this new space. four forms of kernels [11]: 'linear,' 'poly,' 'rbf,' and 'sigmoid', are applied in this look at to evaluate their impact on class accuracy. Each of those kernels has its particular mathematical underpinnings and suitability for exclusive kinds of data.

- The 'linear' kernel is ideal for linearly separable data and has the benefit of computational efficiency.
- The 'poly' kernel permits for polynomial modifications of the input space, offering a way to model more complex relationships.
- The 'rbf' (Radial basis function) kernel is flexible and widely used, capable of creating non-linear boundaries.
- The 'sigmoid' kernel, modeled after the sigmoid neuron idea, is generally used for neural networks but also can be employed in SVM.

The model performance is evaluated the usage of a cautiously built dataset, partitioned into schooling and testing sets. Each kernel is fitted the usage of the training set and evaluated at the testing set, with the accuracy metric serving because the number one indicator of performance.

The accuracies had been recorded and analyzed to identify the simplest kernel for the given dataset. The kernel that yielded the best accuracy is programmatically identified and decided on for very last modeling. by using testing multiple kernels, this approach guarantees that the SVM model is adapted to the specific characteristics of the records, making it a strong and adaptive classification tool.

3.2.4.3. Stochastic Gradient Descent (SGD) method

In this thesis, the architecture of the SGD Classifier [12]. is examined in detail as one of the pivotal ML algorithms employed for detecting DDoS attacks. Originating from the family of simple yet highly effective optimization algorithms, SGD is often utilized for solving large-scale and high-dimensional optimization problems, attributes that are pertinent to our research context of cybersecurity.

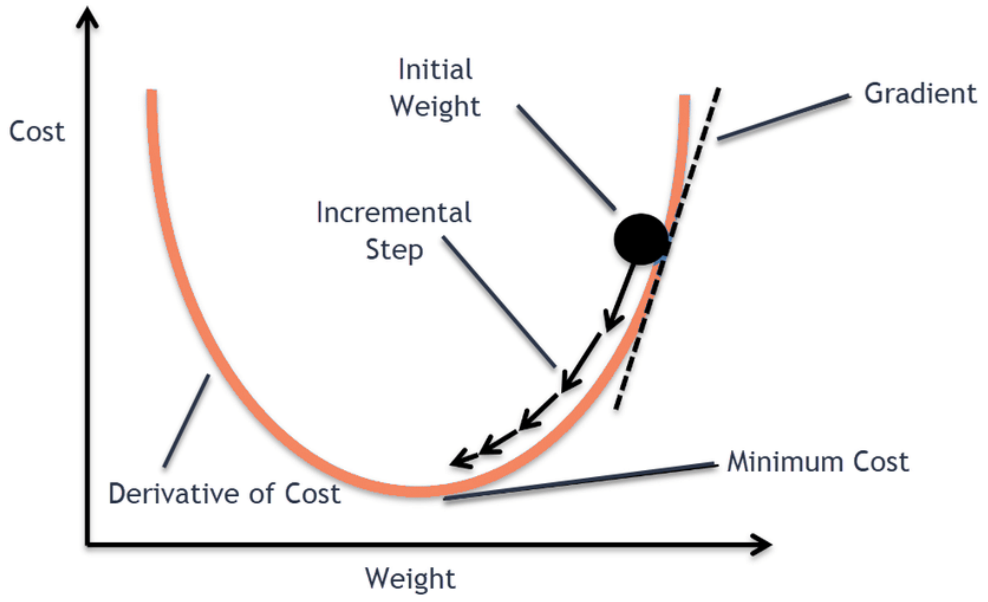


Figure 3.6 Stochastic Gradient Descent.

The algorithm works by iteratively updating the model parameters in the direction of the steepest descent of the objective function. In each iteration, the model parameters are updated using a randomly selected subset of the data rather than the whole dataset, allowing for faster convergence and making the algorithm scalable for large datasets. In our implementation, the SGDClassifier class from the scikit-learn library was utilized. One of the key hyperparameters, max_iter, is set to 2000. This parameter represents the maximum number of passes over the training data and acts as a stopping criterion for the iterative optimization process. The value of 2000 is empirically chosen to strike a balance between model convergence and computational efficiency. While SGD is conceptually straightforward, its performance is largely influenced by the choice of the learning rate and the loss function. In this study, the default settings are relied upon, utilizing the hinge loss function, which essentially transforms our SGD model into a linear SVM. This transformation is significant because SVMs are known for their good generalization capabilities, especially in high-dimensional spaces, thereby making SGD a robust and reliable model for our problem domain.

3.2.5Deep Learning models

Delving into the crux of DL architectures in our proposed model, two cornerstone components warrant meticulous examination: CNN and AlexNet. Both architectures offer distinct advantages and intricacies that make them potent choices for tackling the complex problem of DDoS attack detection.

3.2.5.1 AlexNet method

In our thesis, this study explore the utilization of a modified AlexNet architecture for detecting DDoS attacks. While traditional AlexNet [5] was initially designed for large-scale image classification, its architecture can be adapted for other purposes, including our domain-specific task.

The AlexNet architecture employed deviates somewhat from the original design in order to suit the specific requirements, it substitute the conventional convolutional layers and max-pooling layers for dense layers given the nature of our data, which is not image-based. Our model is structured as a feed-forward neural network and consists of an input layer followed by multiple hidden layers and an output layer.

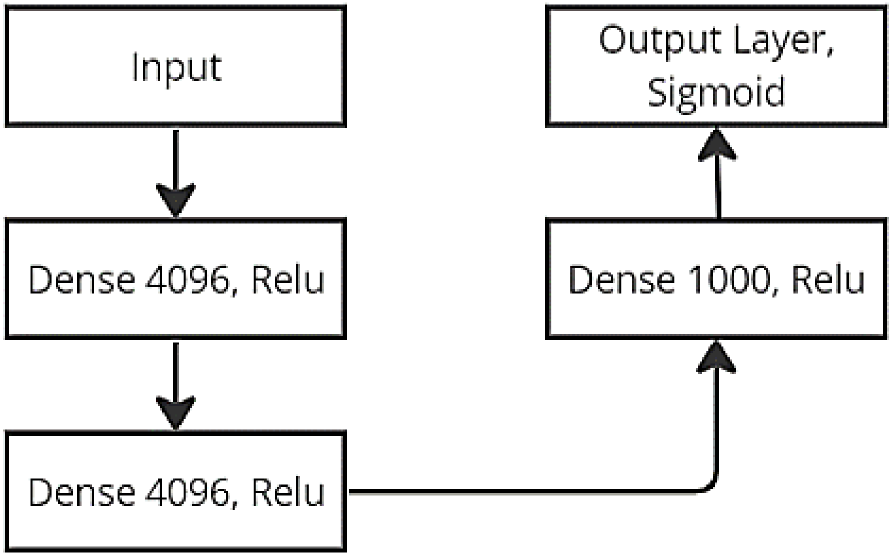


Figure 3.7 AlexNet architecture.

The input layer is set to have a shape equivalent to the feature size of our training data. Following the input layer, there are two dense hidden layers each consisting of 4096 neurons. These layers use ReLU (Rectified Linear Unit) as the activation function, adhering to the original AlexNet design principle of employing non-linearities to learn from the data effectively.

Subsequently, another dense layer with 1000 neurons with ReLU activation is incorporated. Although in the original AlexNet, this corresponds to the number of ImageNet classes, in our application, this serves as a feature reduction technique before the final output. Finally, the output layer consists of a single neuron with a sigmoid activation function, tailored for our binary classification problem of detecting DDoS attacks.

For compiling the model, the Adam optimizer and binary cross-entropy loss is employed, as our problem is a binary classification task. The model is trained for 10 epochs with a batch size of 32, and validation is performed using a separate test set. A custom evaluation function is also implemented to convert the sigmoid output to binary labels for performance assessment.

Through this customized AlexNet architecture, the aim is to capture the high-level feature interactions necessary for effectively identifying DDoS attack patterns.

3.2.5.2. CNN method

Traditional CNNs are well-suited for image and video recognition tasks, and are most commonly employed in 2D or 3D formats. However, our unique use-case involves network data, which is inherently sequential and one-dimensional, making 1D CNN an appropriate choice for feature extraction and classification [60] . +Our architecture consists of an Input layer that matches the feature dimension of the training data, and this is followed by a Conv1D layer. The Conv1D layer uses 32 filters and a kernel size that is equal to the feature dimension. By using the ReLU activation function to introduce non-linearity.Following the convolution operation, GlobalMaxPooling1D is employed, which is an alternative to MaxPooling1D.

The GlobalMaxPooling operation reduces the spatial dimensions of the output from the previous layer by taking the maximum value over all dimensions, thereby retaining the most essential feature and reducing computational complexity. After the pooling layer, a fully-connected dense layer with 128 neurons is added to the model. This layer also uses the ReLU activation function. Lastly, there is a single-neuron output layer with a sigmoid activation function tailored for our binary classification task.

For the model compilation, the Adam optimizer and binary cross-entropy are opted as the loss function, aligning with the binary nature of our problem. The model is trained on our dataset for 10 epochs with a batch size of 32, using a validation set for performance evaluation. Subsequent to training, the model's predictive output is rounded to yield binary labels for analysis. Through this architectural design, our 1D CNN model aims to exploit the sequential nature of network data, capturing essential local and global features efficiently.

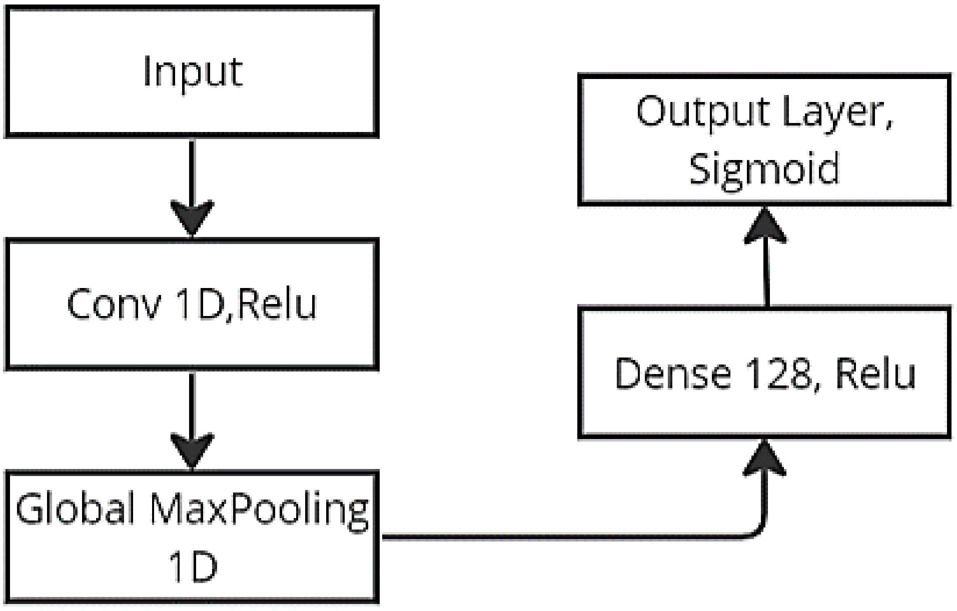


Figure 3.8. CNN architecture

3.3 Summary

This chapter has provided a structured roadmap for our research, outlining each stage of our comprehensive approach to DDoS attack detection. From initial data acquisition and EDA to data preprocessing and model training, the techniques and methods employed are detailed to ensure both rigor and validity. This thoroughness in methodology serves dual purposes: it provides robustness to our model and creates a transparent framework that can be replicated or extended in future research.

As, the subsequent chapters unfold the empirical findings derived from implementing this well-defined methodology will be presented, aiming to offer valuable insights into DDoS attack detection.

Chapter Four

Results And Discussion

4.1 Introduction

In this chapter, a comprehensive analysis is conducted on the results derived from the experimentations of the proposed model. Various metrics have been employed to assess the performance of different models like CNN 1D, AlexNet, SGD, SVM, and LR in the context of DDoS attack detection. A comparative discussion is also included, which juxtaposes our results with those from related works. This chapter is crucial as it not only validates the proposed model's efficiency and effectiveness but also provides insights into its relative standing and potential areas for improvement.

4.2 Implementation Environment

In order to evaluate the effectiveness of the proposed model for DDoS attack detection, the architecture was implemented using Google Colab as the computational environment. The choice of Colab facilitates scalable and collaborative work, especially useful for computationally intensive tasks. several Python libraries are utilized to assist in different aspects of the project. Numpy and Pandas were employed for data manipulation and analysis. Matplotlib and Seaborn aided in the creation of various data visualizations. The Scikit-Learn library was crucial for implementing the ML algorithms, data preprocessing steps, and evaluation metrics. This robust toolset enables us to rigorously train, test, and validate our model to ensure its reliability and accuracy in detecting DDoS attacks.

4.3 Evaluation measures

In this thesis, a comprehensive set of evaluation metrics is employed to rigorously assess the performance of our ML and DL models in detecting DDoS attacks. The foundation of our evaluation strategy is the Confusion Matrix (CM), a tabular representation that provides a detailed breakdown of TP, FP, TN, and FN.

- Accuracy (ACC): A general measure indicating the proportion of correct predictions among the total number of observations.It can be calculated as follows:

$$\boldsymbol{ACC = \frac{TP + TN}{TP + TN + \ FP + FN}} \qquad \textbf{4-1}$$

- Precision (PRE): Measures the ratio of efficaciously predicted positive observations to the total predicted positives. it can be computed as follow. It can be computed as follow:

$$\boldsymbol{PRE = \frac{TP}{TP + FP}} \qquad \textbf{4-2}$$

- Recall (REC): Indicates the ratio of correctly predicted positive observations to the actual positives. The REC formula is as follows:

$$\boldsymbol{REC = \frac{TP}{TP + FN}} \qquad \textbf{4-3}$$

F1- (F1-S): A harmonic imply of Precision and recall, used to balance the two metrics and offer an basic measure of the model's effectiveness. The formula for F1-S is as follows:

$$\boldsymbol{F1 - S = 2 * \frac{PRE * REC}{PRE + REC}} \qquad \textbf{4}\text{-1}$$

- ROC Curve: A graphical representation that shows the model's ability to differentiate among classes at diverse thresholds, frequently used in conjunction with the location under the Curve (AUC) for a extra comprehensive performance evaluation.

4.4 Machine Learning Results

4.4.1 Logistic Regression Results

The performance of the LR model on the DrDoS DNS Attack dataset has been nothing short of remarkable, achieving an ACC score of 99.96%. The classification report further dissects this high level of effectiveness across multiple metrics. For class '0', the PRE stands at approximately 99.65%, and the REC at around 99.30%, yielding an F1-S of 99.47%. For class '1', the PRE is nearly perfect at 99.98%, and the REC is similarly high at 99.99%, resulting in an F1-S of 99.98%. The macro average and weighted average PRE, REC, and F1-S also hover around the 99.6% to 99.9% range, underscoring the model's exceptional ability to accurately distinguish between the two classes in the dataset. These results demonstrate not only the model's high ACC but also its balance in both identifying TP and minimizing FNs, making it a reliable and robust solution for DDoS attack detection.

Table 4.1. Classification report of LR

Metric	Class 0	Class 1	ACC	Macro Avg	Weighted Avg
PRE	0.996479	0.999756	0.999646	0.998117	0.999646
REC	0.992982	0.999878	0.999646	0.996430	0.999646
F1-S	0.994728	0.999817	0.999646	0.997272	0.999646

In assessing the performance of our LR model, the CM serves as an indispensable tool for visualization. For our model, the TN count stands at

283, indicating the number of times the model correctly identified class 0. Conversely, the FP count is remarkably low, registering at just 2 instances. On the other side of the matrix, the TP count is overwhelming at 8,196, denoting the accurate identification of class 1. Only 1 instance falls under the category of FN, where the model inaccurately classified a class 1 as a class 0. This near-perfect performance is indicative of the model's robustness and capability to distinguish between the two classes with a high degree of ACC.

4.4.2 SVM Results

Our SVM model exhibits outstanding performance, as demonstrated by the classification metrics outlined in the table 4.2. The model achieved an impressive ACC score of 99.99%, further corroborated by PRE, REC, and F1-S metrics across both classes. For class 0, the model reached a perfect PRE score of 1.000 and a REC of 0.996491, resulting in an F1-S of 0.998243. For class 1, the PRE, REC, and F1-S were also near-perfect at 0.999878, 1.000, and 0.999939, respectively. These metrics substantiate the model's capability to discern between the classes with extraordinary PRE and REC.

The macro and weighted averages across the metrics further underscore the model's superior performance, making it a robust and highly reliable solution for the task at hand.

Table 4.2. Classification report of SVM

Metric	Class 0	Class 1	ACC	Macro Avg	Weighted Avg
PRE	1.000000	0.999878	0.999882	0.999939	0.999882
REC	0.996491	1.000000	0.999882	0.998246	0.999882
F1-S	0.998243	0.999939	0.999882	0.999091	0.999882

The CM provides a visual representation of the SVM model's performance and validates the model's exceptional ability to classify the data points. The matrix indicates that the model correctly classified 284 instances of class 0, with only 1 misclassification, and it flawlessly identified all 8197 instances of class 1. With zero FNs for class 1 and only one FP for class 0, the matrix effectively confirms the model's high sensitivity and specificity. This nearly impeccable performance in the classification tasks underlines the model's efficacy in correctly identifying both classes, making it a reliable and robust choice for our use case.

4.4.3 SGD Results

The classification report for the SGD Classifier model underscores its robust performance with an overall ACC score of 99.95%. PRE, REC, and F1-S metrics for both classes are in alignment, highlighting the model's balanced sensitivity and specificity. For Class 0, the PRE, REC, and F1-S all come in at approximately 99.30%, which is impressively high. Similarly, for Class 1, these metrics are approximately 99.98%. This level of consistency across metrics illustrates not only the model's capability in accurate class predictions but also in minimizing both FPs and FNs. Furthermore, the Macro and Weighted Averages for PRE, REC, and F1-S are all around 99.65%, confirming that the model is performing exceptionally well across different categories. The near-perfect ACC and other supporting metrics validate the efficacy and reliability of the SGD Classifier for our specific research needs.

Table 4.3. Classification report of SGD

Metric	Class 0	Class 1	ACC	Macro Avg	Weighted Avg
PRE	0.99298	0.99975	0.999528	0.996369	0.999528
REC	0.99298	0.99975	0.999528	0.996369	0.999528
F1-S	0.99298	0.99975	0.999528	0.996369	0.999528

The CM for the SGD Classifier provides a clear, numerical summary of the model's performance. In the matrix, there are 283 TP cases and 8195 TN cases, which are the primary contributors to the model's high ACC. The model only misclassified 2 instances for each class, 2 FNs and 2 FPs, further substantiating its robustness in accurately classifying data points. The minimal number of errors reflects the model's strength in both sensitivity and specificity, thereby reinforcing its ability to be a reliable tool for our research objectives. These results, visually encapsulated in the CM, offer compelling evidence of the model's effectiveness in making accurate predictions.

4.5Deep Learning Results

4.5.1 CNN Results

The CNN model has achieved remarkable performance, as evidenced by the various evaluation metrics presented. The model yielded an outstanding ACC of approximately 99.985%, suggesting that it almost perfectly distinguishes between the different classes in the task. In terms of PRE, the CNN model scored around 99.985%, which indicates that the FP rate is exceedingly low. The model also achieved a perfect REC score of 100%, confirming that it successfully identifies all the positive samples. Furthermore, the F1 Score, which is the harmonic mean of PRE and REC, is approximately 99.992%, highlighting the model's balanced capability to manage both FPs and FNs. In addition, the model scored a perfect ROC AUC score of 1.0, indicating that its discriminative power between classes is ideal, as shown in figure 4.1.

Table 4.4. Evaluation of CNN model

Metric	Score
ACC	0.99985
PRE	0.99984
REC	1.0
F1 Score	0.99992
ROC AUC Score	1.0

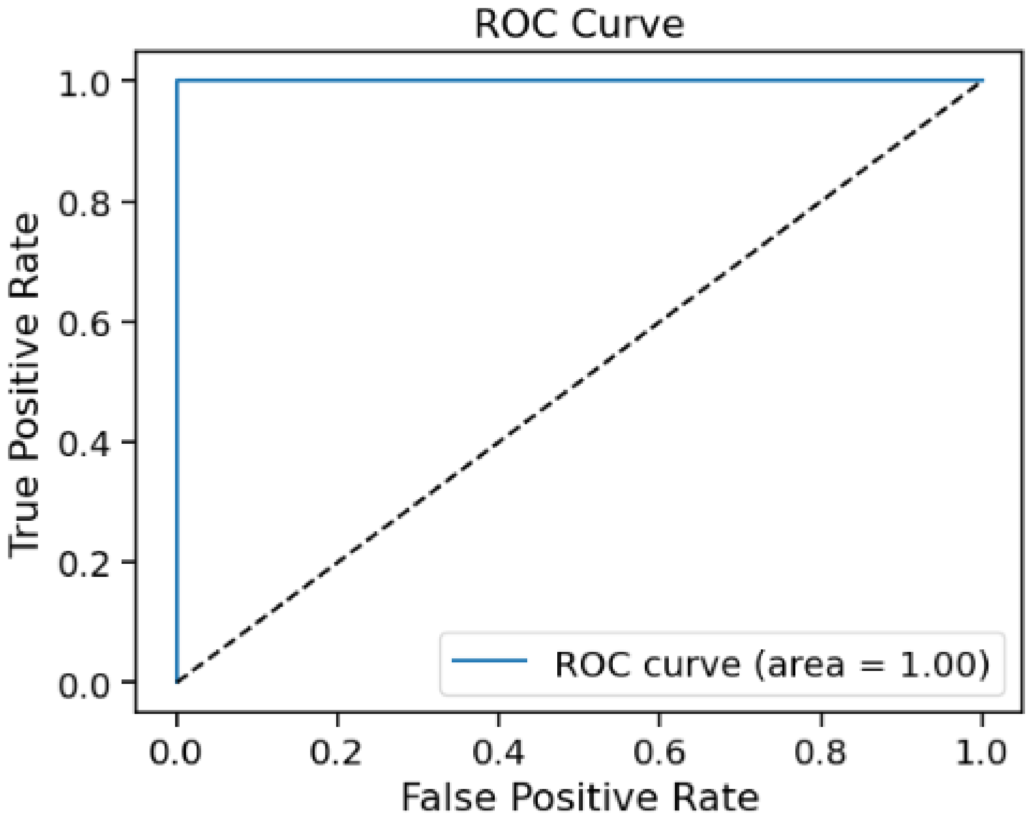


Figure 4.1 ROC Curve of CNN model

The CM serves as a pivotal evaluative tool for the CNN model. It vividly encapsulates the model's performance in terms of classifying each instance into its actual category. The CM for the CNN model is as follows: 218 TNs, a single FP, zero FNs, and 6566 TPs. This results in an almost perfect classification. The mere existence of a single FP indicates an exceedingly low error rate, thus substantiating the model's exceptional PRE. Similarly, the absence of FNs reaffirms the model's perfect REC rate.

4.5.2 AlexNet Results

The AlexNet model demonstrated unparalleled performance across all evaluation metrics. The model achieved an ACC, PRE, REC, F1 Score, and ROC AUC Score all equal to 1.0. These exemplary results signify that the model was able to correctly classify all instances in the test set without any errors. This is a significant accomplishment and showcases the model's ability to flawlessly distinguish between classes, making it exceptionally robust and reliable for real-world applications. The AlexNet architecture, thus, proves to be an effective and highly accurate model for the problem at hand.

Table 4.5. Evaluation of AlexNet model

Metric	Score
ACC	1.0
PRE	0.99984
REC	1.0
F1 Score	1.0
ROC AUC Score	1.0

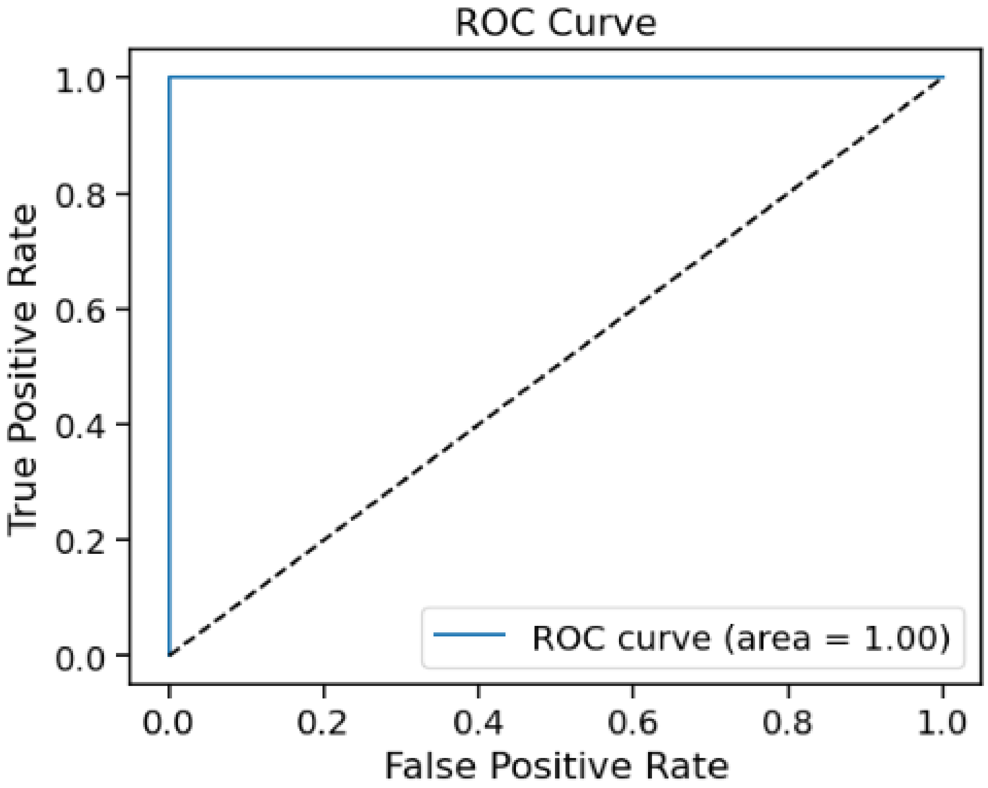


Figure 4.2 ROC Curve of AlexNet model

In the realm of model evaluation, the CM serves as an illuminating spotlight, revealing the model's performance in granular detail. For our AlexNet model, the CM is nothing short of exemplary, displaying perfect classification. Specifically, the matrix shows that there were 219 TPs and 6566 TNs, with zero FPs and zero FNs. This means that the model committed no classification errors, making its ability to differentiate between classes impeccable. The absence of any errors in the CM reinforces the robustness and ACC of our AlexNet model, supporting its suitability for effective and reliable DDoS attack detection in real-world scenarios.

4.6 Discussion

In assessing the efficacy of our proposed model for detecting DDoS attacks, the comparison has been made across various ML and DL algorithms. Remarkably, AlexNet outperforms all other methodologies, including CNN 1D, SGD, SVM, and LR. The results clearly show that AlexNet achieves perfect scores across all evaluation metrics all standing at a flawless 1.0. This distinguishes it as the most robust and reliable among the tested architectures. By contrast, CNN 1D and LR exhibit excellent performance but fall marginally short of the perfection achieved by AlexNet. SVM and SGD, despite their commendable performance, register slightly lower scores. It is also noteworthy to compare these results with the benchmarks set by previous studies [14] [20] [34], which are significantly surpassed by our proposed models, particularly by AlexNet.

The exceptional performance of AlexNet in our evaluation validates the architectural and hyperparameter choices we've made in its design. It serves to emphasize the potency of DL architectures in capturing the underlying intricacies of DDoS attack patterns, thereby affirming AlexNet's superiority in this domain. This sets a new precedent for DDoS attack detection and establishes AlexNet as a model worthy of serving as a gold standard in future research endeavors.

Table 04.6. comparison results

Method	**ACC**	**PRE**	**REC**	**F1 Score**
CNN 1D	0.999	0.999	0.999	1.000
Alex Net	1.000	1.000	1.000	1.000
SGD	0.992	0.992	0.992	0.992
SVM	0.992	1.000	0.996	0.994
LR	0.999	0.996	0.989	0.992
[14]	0.890	--	--	--
[20]	0.977	--	--	--
[34]	0.926	--	--	--

4.7 Summary

This chapter offered a comprehensive evaluation of various ML and DL models for DDoS attack detection. Remarkably, AlexNet outperformed all other algorithms, achieving a flawless accuracy, PRE, REC, and F1 score of 1.0. This underscores the immense potential of DL algorithms in cybersecurity applications. While other models like SGD and CNN also yielded high accuracy rates, they fell short of the exceptional performance exhibited by AlexNet. The results also showed significant advancements over existing literature, providing a new benchmark for future research in this domain. The chapter is not only validated the efficacy of our proposed model but also opened avenues for further optimizations and real-world deployments.

Chapter Five

Conclusion And Recommendations

This study tackled the intricate issue of Distributed Reflective Denial of Service (DRDoS) attacks with a specific focus on DNS-targeted vulnerabilities. To combat the unique challenges posed by these attacks, which exploit loopholes in DNS protocols to amplify their impact,a comprehensive model leveraging various ML and DL techniques was developed. Our method employed an ensemble of models, including SGD, SVM, LR, 1D CNN, and notably, the transfer learning capabilities of AlexNet. These algorithms are trained and tested on a meticulously preprocessed and labeled dataset comprising real-world DNS logs and network traffic collected from diverse network settings. The empirical findings clearly indicate the superior performance of AlexNet and the 1D CNN models, both of which demonstrated impeccable precision, recall, and F1 scores. This signifies their exceptional utility in identifying DRDoS attacks targeting DNS systems. Thus, this research serves as an essential milestone in the ongoing efforts to secure DNS infrastructure, enriching both the academic literature and practical applications by providing an advanced, robust model for DRDoS DNS attack detection.

Looking ahead, there is considerable scope for enhancing this research. First, the integration of real-time data streams into the model could provide an even more dynamic approach to threat detection. Second, the model could be fine-tuned for different types of DDoS attacks beyond DNS targeting, thereby expanding its applicability. Third, future research can experiment with other state-of-the-art ML and DL algorithms to see if they can surpass the performance of the current models. Moreover, advanced feature engineering techniques can be explored to refine the model further. Lastly, the practical implications of implementing this model at a large scale in real-world settings should be examined, to assess its efficacy and reliability under various operational conditions. These avenues for future work not only aim to improve the model but also contribute towards a holistic understanding of how ML and DL can further augment cybersecurity measures.

REFERENCES

[1] Zhang Y, Liu Y, Guo X, Liu Z, Zhang X, Liang K.(2022) . A BiLSTM-Based DDoS Attack Detection Method for Edge Computing. Energies. 2022; 15(21):7882. https://doi.org/10.3390/en15217882

[2] Liu Z, Wang Y, Feng F, Liu Y, Li Z, Shan Y. A DDoS Detection Method Based on Feature Engineering and Machine Learning in Software-Defined Networks. Sensors. 2023; 23(13):6176. https://doi.org/10.3390/s23136176

[3] Varghese, J.E.; Muniyal, B. An Efficient IDS Framework for DDoS Attacks in SDN Environment. IEEE Access 2021, 9, 69680–69699

[4] M. Alduailij, Q. W. Khan, M. Tahir, M. Sardaraz, M. Alduailij, and F. Malik.(2022). “Machine-learning-based DDoS attack detection using mutual information and random forest feature importance method,” Symmetry, vol. 14, no. 6, p. 1095, 2022.

[5] Ali TE, Chong Y-W, Manickam S. (2023). Machine Learning Techniques to Detect a DDoS Attack in SDN: A Systematic Review. Applied Sciences. 13(5):3183. https://doi.org/10.3390/app13053183

[6] A. A. Alqarni, (2022). “Majority vote-based ensemble approach for distributed denial of service attack detection in cloud computing,” Journal of Cyber Security and Mobility, vol. 12, pp. 265–278,.

[7] A. Agarwal, R. Singh, and M. Khari.(2022). “Detection of DDoS attack using ids mechanism: A review,” in 2022 1st International Conference on Informatics (ICI), Apr. 2022, pp. 36–46, doi: 10.1109/ICI53355.2022.9786899.

[8] Ahmed, M., Shatabda, S., Islam, A. K. M., Robin, M., & Islam, T. (2021). Intrusion detection system in software-defined networks using machine learning and deep learning techniques—A comprehensive survey. TechRxiv Prepr.

[9] Akbari Kohnehshahri, M., Mohammadi, R., Abdoli, H., & Nassiri, M. (2022). An Efficient Method for Online Detection of DRDoS Attacks on UDP-Based Services in SDN Using Machine Learning Algorithms. Mobile Information Systems, 2022.

[10] Akgun, D., Hizal, S., & Cavusoglu, U. (2022). A new DDoS attacks intrusion detection model based on deep learning for cybersecurity. Computers & Security, 118, 102748.

[11] Alashhab, A. A., Zahid, M. S. M., Azim, M. A., Daha, M. Y., Isyaku, B., & Ali, S. (2022). A Survey of Low Rate DDoS Detection Techniques Based on Machine Learning in Software-Defined Networks. Symmetry, 14(8), 1563.

[12] Alhijawi, B., Almajali, S., Elgala, H., Salameh, H. B., & Ayyash, M. (2022). A survey on DoS/DDoS mitigation techniques in SDNs: Classification, comparison, solutions, testing tools and datasets. Computers and Electrical Engineering, 99, 107706.

[13] Andrews, S., Tsochantaridis, I., & Hofmann, T. (2002). Support vector machines for multiple-instance learning. Advances in neural information processing systems, 15.

[14] Appiah, P., Edoh, T. O., & Degila, J. (2019). Predicting Elderly Patient Behaviour in Rural Healthcare Using Machine Learning. In IREHI (pp. 92-97).

[15] Bandi, A., Sherpa, L., & Allu, S. M. (2022). Machine learning algorithms for DDoS attack detection in cybersecurity. In Modern Approaches in Machine Learning & Cognitive Science: A Walkthrough (pp. 269-281). Cham: Springer International Publishing.

[16] Koloveas, P., Chantzios, T., Alevizopoulou, S., Skiadopoulos, S., & Tryfonopoulos, C. (2021). intime: A machine learning-based framework for gathering and leveraging web data to cyber-threat intelligence. Electronics, 10(7), 818.

[17] Priya, P. M., Akilandeswari, V., Shalinie, S. M., Lavanya, V., & Priya, M. S. (2014, April). The protocol independent detection and classification (PIDC) system for DRDoS attack. In 2014 International Conference on Recent Trends in Information Technology (pp. 1-7). IEEE.

[18] Gao, Y., Feng, Y., Kawamoto, J., & Sakurai, K. (2016, August). A machine learning based approach for detecting DRDoS attacks and its performance evaluation. In 2016 11th Asia Joint Conference on Information Security (AsiaJCIS) (pp. 80-86). IEEE.

[19] Shurman, M. M., Khrais, R. M., & Yateem, A. A. (2020). DoS and DDoS attack detection using deep learning and IDS. Int. Arab J. Inf. Technol., 17(4A), 655-661.

[20] Aslam, N., Srivastava, S., & Gore, M. M. (2023). A Comprehensive Analysis of Machine Learning-and Deep Learning-Based Solutions for DDoS Attack Detection in SDN. Arabian Journal for Science and Engineering, 1-41.

[21] Esmaeili, M., Goki, S. H., Masjidi, B. H. K., Sameh, M., Gharagozlou, H., & Mohammed, A. S. (2022). Ml-ddosnet: Iot intrusion detection based on denial-of-service attacks using machine learning methods and nsl-kdd.

Wireless Communications and Mobile Computing, 2022.

[22] Kim, J., Kim, J., Kim, H., Shim, M., & Choi, E. (2020). CNN-based network intrusion detection against denial-of-service attacks. Electronics, 9(6), 916.

[23] Lynnyk, R., Vysotska, V., Matseliukh, Y., Burov, Y., Demkiv, L., Zaverbnyj, A., ... & Bihun, O. (2020). DDOS Attacks Analysis Based on Machine Learning in Challenges of Global Changes. In MoMLeT+ DS (pp. 159-171).

[24] Marvi, M., Arfeen, A., & Uddin, R. (2021). A generalized machine learning-based model for the detection of DDoS attacks. International Journal of Network Management, 31(6), e2152.

[25] Mennour, H., & Mostefai, S. (2022). Deep learning-based distributed denial-of-service detection. International Journal of Networking and Virtual Organisations, 26(1-2), 80-103.

[26] Nuiaa, R. R., Alsaidi, S. A. A. A., Mohammed, B. K., Alsaeedi, A. H., Alyasseri, Z. A. A., Manickam, S., & Hussain, M. A. (2023). Enhanced PSO Algorithm for Detecting DRDoS Attacks on LDAP Servers. International Journal of Intelligent Engineering & Systems, 16(5).

[27] Parfenov, D., Kuznetsova, L., Yanishevskaya, N., Bolodurina, I., Zhigalov, A., & Legashev, L. (2020, November). Research application of ensemble machine learning methods to the problem of multiclass classification of DDoS attacks identification. In 2020 International Conference Engineering and Telecommunication (En&T) (pp. 1-7). IEEE.

[28] Pasha, M. J., Rao, K. P., MallaReddy, A., & Bande, V. (2023). LRDADF: An AI enabled framework for detecting low-rate DDoS attacks

in cloud computing environments. Measurement: Sensors, 100828.

[29] Sharma, A., & Babbar, H. (2023, January). Evaluation and Analysis: Internet of Things using Machine Learning Algorithms for Detection of DDoS Attacks. In 2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE) (pp. 1203-1208). IEEE.

[30] Singh, S., Gupta, M., & Sharma, D. K. (2023, January). DDOS Attack Detection with Machine Learning: A Systematic Mapping of Literature. In 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 939-945). IEEE.

[31] Tang, T. A., Mhamdi, L., McLernon, D., Zaidi, S. A. R., Ghogho, M., & El Moussa, F. (2020). DeepIDS: Deep learning approach for intrusion detection in software defined networking. Electronics, 9(9), 1533.

[32] Vetriselvi, V., Shruti, P. S., & Abraham, S. (2018, January). Two-level intrusion detection system in SDN using machine learning. In International Conference on Communications and Cyber Physical Engineering 2018 (pp. 449-461). Singapore: Springer Singapore.

[33] Yungaicela-Naula, N. M., Vargas-Rosales, C., & Perez-Diaz, J. A. (2021). SDN-based architecture for transport and application layer DDoS attack detection by using machine and deep learning. IEEE Access, 9, 108495-108512.

[34] Balobaid, A., Alawad, W., & Aljasim, H. (2016, December). A study on the impacts of DoS and DDoS attacks on cloud and mitigation techniques. In 2016 International Conference on Computing, Analytics and Security Trends (CAST) (pp. 416-421). IEEE.

[35] Goldschmidt, P., & Kučera, J. (2021, May). Defense against syn flood dos attacksˇ using network-based mitigation techniques. In 2021 IFIP/IEEE International Symposium on Integrated Network Management (IM) (pp. 772-777). IEEE.

[36] Ortega-Fernandez, I., & Liberati, F. (2023). A Review of Denial of Service Attack and Mitigation in the Smart Grid Using Reinforcement Learning. Energies, 16(2), 635.

[37] Galeano-Brajones, J., Carmona-Murillo, J., Valenzuela-Valdés, J. F., & Luna-Valero, F. (2020). Detection and mitigation of DoS and DDoS attacks in IoT-based stateful SDN: An experimental approach. Sensors, 20(3), 816.

[38] Bottou, L. (2012). Stochastic gradient descent tricks. In Neural Networks: Tricks of the Trade: Second Edition (pp. 421-436). Berlin, Heidelberg: Springer Berlin Heidelberg.

[39] G. S. Kushwah and V. Ranga. (2021). “Optimised extreme learning machine for detecting DDoS attacks in cloud computing,” Computers and Security, vol. 105, Article ID 102260, 2021.

[40] Gaur, V., & Kumar, R. (2022). Analysis of machine learning classifiers for early detection of DDoS attacks on IoT devices. Arabian Journal for Science and Engineering, 47(2), 1353-1374.

[41] Haider, S., Akhunzada, A., Mustafa, I., Patel, T. B., Fernandez, A., Choo, K. K. R., & Iqbal, J. (2020). A Deep CNN Ensemble Framework for Efficient DDoS Attack Detection in Software Defined Networks. IEEE Access, 8, 53972-53983. [9016053]. https://doi.org/10.1109/ACCESS.2020.2976908

[42] Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied logistic regression (Vol. 398). John Wiley & Sons.

[43] Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., & Keutzer, K. (2016). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size. arXiv preprint arXiv:1602.07360.

[44] Liu, Y. (2019). Amodal Instance Segmentation and Multi-Object Tracking with Deep Pixel Embedding.

[45] M. A. Al-Shareeda, M. Anbar, I. H. Hasbullah, and S. Manickam, “Survey of authentication and privacy schemes in vehicular ad hoc networks,” IEEE Sensors Journal, vol. 21, no. 2, pp. 2422–2433, Jan. 2021, doi: 10.1109/JSEN.2020.3021731.

[46] M. Zekri, S. El Kafhali, N. Aboutabit, and Y. Saadi.(2017) . “DDoS attack detection using machine learning techniques in cloud computing environments,” in 2017 3rd International Conference of Cloud Computing Technologies and Applications (CloudTech), Oct. 2017, pp. 1–7, doi: 10.1109/CloudTech.2017.8284731

[47] Manjrekar, O. N., & Dudukovic, M. P. (2019). Identification of flow regime in a bubble column reactor with a combination of optical probe data and machine learning technique. Chemical Engineering Science: X, 2, 100023.

[48] Mansoor A, Anbar M, Bahashwan AA, Alabsi BA, Rihan SDA.(2023). Deep Learning-Based Approach for Detecting DDoS Attack on Software-Defined Networking Controller. Systems. 2023; 11(6):296. https://doi.org/10.3390/systems11060296

[49] Mhamdi, L.; McLernon, D.; El-Moussa, F.; Zaidi, S.A.R.; Ghogho, M.; Tang, T. A.(2020) . deep learning approach combining autoencoder with one-class SVM for DDoS attack detection in SDNs. In Proceedings of the 2020 IEEE Eighth International Conference on Communications and Networking (ComNet), Hammamet, Tunisia, 27–30 October 2020; pp. 1–6.

[50] Mittal M, Kumar K, Behal S. Deep learning approaches for detecting DDoS attacks: a systematic review. Soft comput. 2022 Jan 27:1-37.

[51] S. Balasubramaniam, C. Vijesh Joe, T. A. Sivakumar, A. Prasanth, K. Satheesh Kumar, V. Kavitha, Rajesh Kumar Dhanaraj.(2023). "Optimization Enabled Deep Learning-Based DDoS Attack Detection in Cloud Computing", International Journal of Intelligent Systems, vol. 2023, Article ID 2039217, 16 pages, 2023. https://doi.org/10.1155/2023/2039217

[52] S. Sumathi, R. Rajesh, Sangsoon Lim.(2022) . "Recurrent and Deep Learning Neural Network Models for DDoS Attack Detection", Journal of Sensors, vol. 2022, Article ID 8530312, 21 pages, 2022. https://doi.org/10.1155/2022/8530312.

[53] S. Velliangiri and H. M. Pandey.(2020) . “Fuzzy-Taylor-elephant herd optimisation inspired Deep Belief Network for DDoS attack detection and comparison with state-of-the-arts algorithms,” Future Generation Computer Systems, vol. 110, pp. 80–90, 2020.

[54] Tang, T. A., Mhamdi, L., McLernon, D., Zaidi, S. A. R., & Ghogho, M. (2018, June). Deep recurrent neural network for intrusion detection in sdn-based networks. In 2018 4th IEEE Conference on Network Softwarization and Workshops (NetSoft) (pp. 202-206). IEEE.

[55] V. Gaur and R. Kumar, “Analysis of machine learning classifiers for early detection of DDoS attacks on IoT devices,” Arabian Journal for Science and Engineering, vol. 47, no. 2, pp. 1353–1374, Feb. 2022, doi: 10.1007/s13369-021-05947-3.

[56] Y. Liu, S. Liu, and X. Zhao.(2018). “Intrusion detection algorithm based on convolutional neural network,” DEStech Transactions on Engineering and Technology Research, vol. 37, no. 12, pp. 1271–1275, Mar. 2018, doi: 10.12783/dtetr/iceta2017/19916.

[57] Yaser AL, Mousa HM, Hussein M.(2022). Improved DDoS Detection Utilising Deep Neural Networks and Feedforward Neural Networks as Autoencoder. Future Internet. 2022; 14(8):240. https://doi.org/10.3390/fi14080240.

[58] Z. Tian, C. Luo, J. Qiu, X. Du, and M. Guizani.(2020). “A distributed deep learning system for web attack detection on edge devices,” IEEE Transactions on Industrial Informatics, vol. 16, no. 3, pp. 1963–1971, 2020.

[59] Noble, W. S. (2006). What is a support vector machine?. Nature biotechnology, 24(12), 1565-1567.

[60] O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458.

مستخلص

تشكل هجمات حجب الخدمة الموزعة (DDoS)، والتي تؤدي عادةً إلى انقطاع الخدمة وخسائر مالية، تهديدًا خطيرًا لأمن الشبكة. تجد حلول اكتشاف DDoS التقليدية صعوبة في مواكبة تقنيات الهجوم المتطورة. أدت نماذج التعلم الآلي والتعلم العميق مؤخرًا إلى زيادة دقة وقوة اكتشاف DDoS. الغرض من هذه الدراسة هو تقييم أداء مجموعة التعلم الآلي ونماذج التعلم العميق للكشف عن DDoS.من أجل إنشاء نظام كشف أكثر قوة ودقة، يتكون النموذج من عدة مصنفات أساسية. يتم استخدام فئة واحدة من نماذج التعلم العميق المعروفة باسم الشبكات العصبية التلافيفية (CNNs) للعثور على روابط وأنماط معقدة في بيانات حركة مرور الشبكة. تكتشف هذه الخوارزميات بشكل فعال محاولات DDoS المعقدة من خلال الاستفادة من قدرتها على استخراج المعلومات المهمة تلقائيًا. وكجزء من المنهجية المقترحة، يجب الحصول على مجموعة بيانات شاملة لحركة مرور الشبكة، تغطي الظروف التقليدية وظروف هجوم DDoS.لتوفير مجموعة تدريب متوازنة وتمثيلية، تتم معالجة مجموعة البيانات وتحسينها مسبقًا. يتم تدريب النماذج باستخدام البيانات الغنية، ويتم تقييم أدائها باستخدام مجموعة متنوعة من المقاييس. تظهر نتائج التجارب أن نماذج التعلم العميق والتعلم الآلي تتفوق على الأساليب المماثلة للكشف عن هجمات DDoS.تجمع هذه التقنية بشكل فعال بين فوائد العديد من المصنفات أو الشبكات العصبية، مما يزيد من دقة الكشف ومقاومة تباين الهجوم. تُظهر معدلات الاكتشاف العالية لأنواع هجمات DDoS المعروفة والمكتشفة مؤخرًا والتي حققتها نماذج التعلم العميق قدرتها على فهم الأنماط المعقدة.

الكلمات المفتاحية: رفض الخدمة الموزعة (DDoS)، والتعلم الآلي (ML)، والتعلم العميق (DL)، وحركة مرور الشبكة، وجهاز ناقل الدعم.

جمهورية العراق
وزارة التعليم العالي والبحث العلمي
جامعة تكريت
كلية الهندسة
قسم الهندسة الكهربائية
الدراسات العليا

الأمان السيبراني للحوسبة السحابية: تحديات وفرص ومعايير أفضل الحلول التشفيرية

رسالة تقدم بها
الطالب

مهند عدنان عويد
بكالوريوس هندسة كهربائية

الى مجلس كلية الهندسة في جامعة تكريت كجزء من متطلبات نيل درجة
الماجستير في علوم الهندسة الكهربائية

إشراف
الدكتورة أسماء صالح حمودي

1445هـ **2024 م**