Adaptive Machine Learning Techniques for Real-Time DDoS Attack Identification and Mitigation

By
Md Nahid Hasan Chowdhury
212-15-4125

&

Sadman Khan
212-15-4128

FINAL YEAR DESIGN PROJECT REPORT

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

Supervised by

Mr. Shah Md Tanvir Siddiquee
Assistant Professor
Department of Computer Science and Engineering
Daffodil International University

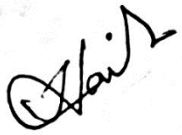
DAFFODIL INTERNATIONAL UNIVERSITY
Dhaka, Bangladesh

May 14, 2025

APPROVAL

This Project titled **"Adaptive Machine Learning Techniques for Real-Time DDoS Attack Identification and Mitigation"**, submitted by Md Nahid Hasan Chowdhury, ID No: 212-15-4125 and Sadman Khan, ID No: 212-15-4128 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **14 May, 2025**.

BOARD OF EXAMINERS



Dr. Sheak Rashed Haider Noori — **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Mohammad Monirul Islam — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

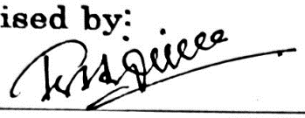
Ms. Tasfia Anika Bushra — **Internal Examiner**
Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Md. Zulfiker Mahmud — **External Examiner**
Professor
Department of Computer Science and Engineering
Jagannath University

DECLARATION

We hereby declare that this project has been done by us under the supervision of **Mr. Shah Md Tanvir Siddiquee**, **Assistant Professor**, Department of Computer Science and Engineering, Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:



Mr. Shah Md Tanvir Siddiquee
Assistant Professor
Department of Computer Science and EngineeringDaffodil International University

Submitted by:

Md Nahid Hasan Chowdhury
Student ID: 212-15-4125
Department of Computer Science and Engineering
Daffodil International University

Sadman Khan
Student ID: 212-15-4128
Department of Computer Science and Engineering
Daffodil International University

ACKNOWLEDGEMENTS

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

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

We are grateful and wish our profound indebtedness to **Mr. Shah Md Tanvir Siddiquee**, **Assistant Professor**, Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh. Deep knowledge and keen interest of our supervisor in the field of **Cloud Computing & Security** to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

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

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

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

ABSTRACT

DDoS or Distributed Denial of Service attacks are primary concern for network security, inundating victims with botnet-generated traffic to render them largely inoperable. This work explores how to utilize machine learning techniques to detect DDoS effectively, using a feature set of various network traffic features and an extracted basic dataset that labels them as either benign or DDoS. In this regard, the research method consisted of strict preprocessing of the data, where it was necessary to eliminate missing values and conduct a train-test split with 70% training data and 30% testing data. This last adjustment was decisive to guarantee realism in the evaluation of the models. A total of five separate machine learning models were implemented; Random Forest, Logistic Regression, Neural Network, SVM or Support Vector Machine, and KNN or K-Nearest Neighbors. A mixed model also was developed using soft voting classifier which combines vote from these models to blend the advantages. Robust evaluation of the models on several metrics was conducted, including accuracy, precision, recall, F1 score, and Area under ROC curve (AUC) score. Confusion matrices and ROC curves gave detailed demonstrations on how well the classification performed, and the models held strong results in classifying benign from malicious traffic. The outcomes indicated that Random Forest model due to its response to complex interaction of features and hybrid due to assembling performed better in detecting. The results show that a combination of different ML techniques improve the DDoS detection accuracy and robustness. This research represents significant contributions towards improving detection mechanisms that can affect greatly the development of more secure networks, including those preventing the impact of DDoS as computational infrastructures become more sophisticated and overtly susceptible to having their operation disrupted by such attacks.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

DDoS or Distributed Denial of Service is a dangerous threat that can shut down any online digital services by bombarding the platforms with so much traffic they become unavailable, leading to financial harm and reputational damage. As more devices connect to the Internet, and cloud storage become common, these attacks are also on the rise and becoming more sophisticated. Traditional detection approaches, such as configuring the traffic threshold, tend to be outdated and stand idle against new types of attacks. Static rules or signature-based approaches have proven ineffective against the evolving tactics behind modern DDoS attacks, and they are also being relied on. It is one of the most dangerous threats to the network. it inundates the system with damaging traffic.

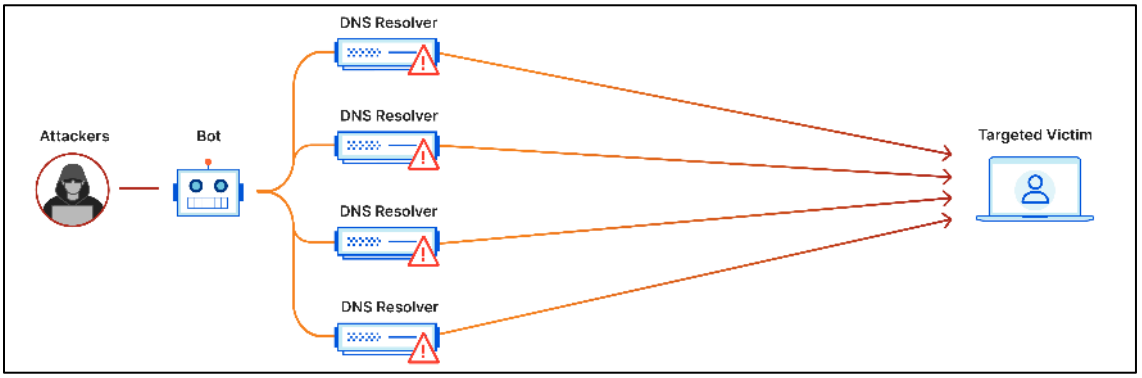


Figure 1.1: DDoS Attack

On the upper figure 1.1, it shows the DDoS concept using graphical object where an attacker in attacking a targeted source using bots which is basically a DDoS attack. So, this thesis shows an efficient way to detect and prevent Data Breach using ML algorithms. To address the limitations of existing approaches, as well as to increase the accuracy and flexibility, this thesis will also propose to include a hybrid machine learning to the current strategies, which entails the combination of diverse techniques to timely detect and classify DDoS attack traffic. Using a hybrid algorithm that combines multiple machine learning models, the system will examine historical network traffic data to spot trends that could point to the potential for DDoS attacks, classify the types of these attacks, and deliver real- time response.

1.2 Motivation

DDoS attacks are more prevalent than ever, threatening to break down the digital infrastructure fueling much of contemporary life and disrupting the services on which businesses, governments and people rely. As businesses have moved to the cloud, increasingly handing sensitive and confidential user data, the size and scale of these attacks have become unprecedented, to the point where they can easily disrupt critical systems and leak invaluable data. The speedy growth of poorly secured IoT or Internet

of Things devices has elevated the surface area for criminal exploitation, making it easier than ever to disrupt infrastructure on a massive scale. DDoS attacks are constantly evolving, now being more innovative and random, traditional detection mechanisms are no longer able to cope with the modern DDoS attacks. This immediate need to ensure availability of cloud-based services are available and sensitive data is secure, highlights the necessity for advanced, adaptive detection and mitigation systems. These threats are not just nicking at a sore spot, they are getting larger, more complex, and more sophisticated than ever before—and it requires more than just a response; we need to be preventing them—and what better way to do so than by combining the latest technology with creative ingenuity to outsmart and better prepare the surface next time? This timely intervention enables us to safeguard our critical systems, restore trust in digital landscapes, and prepare for a safer, more interconnected future. Now is the opportunity for strong solutions to not only combat the growing DDoS threat landscape, but catalyst organizations and users alike to embrace online life in a more connected world.

1.3 Objectives

This thesis aims to:

- Developing an end-to-end model from data collection, feature extraction, training to online classification.
- Assess the performance of various classifiers and find the best approach for DDoS detection.
- Develop a hybrid ML algorithm to detect attacks with a high detection and false positive rate.
- Establish redundancy of data or other inconsistencies in order to correct them from the historical data set which can impact the performance of the trained model.
- Extracting the efficient and important features that will affect the detection accuracy most.
- Overcome issues of data imbalance and computation efficiency in utilizing DDoS detection.

1.4 Methodology

The resulting methodology of this study is aimed at providing a rigorous process that facilitates the evaluation of different machine learning models effective in identifying DDoS attacks using a refined dataset containing labelled features present in a network traffic flow (either as benign or DDoS). This first step constituted of a thorough data prepossessing process, ensuring that the dataset was fit for use for training the model. At first, we performed EDA or Exploratory data analysis to find missing values, and the distribution of the missing values of some selected features was visualized by using histograms and bar plots. To avoid any potential biases in model training, we decided to remove these missing values and have a clean dataset. It also removed the columns related to the infinite values which were replaced with the NaN, but made sure of numerical stability, and turned the target variable 'Label' into a numerical representation; mapping 'BENIGN' as 0 and 'DDoS' as 1 for the purposes of binary classification. For model evaluation, the dataset was divided into 70% for training and 30% for testing datasets (with fixed random state to obtain reproducible results).

We chose a variety of machine learning algorithms to provide a complete evaluation of methods for DDoS detection: Random Forest (RF), Logistic Regression (LR), Neural Network (NN), Linear Support Vector Machine (SVM), and K-Nearest Neighbors (KNN). To improve detection robustness, a hybrid model was created that included the predictions of RF, LR, NN and KNN using soft voting classifier to take advantage of their complementarity. Residual missing values while training the model were also handled with pipelines including for LR, NN, SVM, and KNN using mean imputation to allow for a consistent data treatment. Feature importance was additionally computed and visualized through horizontal bar plots using the RF model, and illustrated the level of discriminative power of network traffic features.

All the models were trained using the training set and tested on the testing set with a strong variety of metrics, which were calculated separately for each model based on the number of true positives, the number of true negatives, as well as the number of false positives and false negatives. Receiver Operating Characteristic (ROC) curves were plotted and the area under the curve (AUC) values provided a numerical value of the models' discriminative power between benign and DDoS traffic. The bar charts were employed to showcase a graphical representation of model performance over all the metrics, and a table was created to quantify the results. The evaluation process also included a large ensemble technique for model evaluation to ensure that both individual model efficacies as well as ensemble method multiresolution were well-addressed for DDoS detection, begetting reliable security solutions.

1.5 Project Outcome

The results obtained by this research paper are very crucial in leveraging machine learning against DDoS attack detection which can then be used to improve network security. We compare five individual models against a hybrid soft voting classifier to verify the effectiveness of combined results in classifying benign and DDoS traffic. Our Random Forest model also performed well with high accuracy, F1 score, and precision, recall and AUC metrics as this model is able to capture more complex feature interactions, and is less susceptible to overfitting. The hybrid model which was based on ensemble learning enhanced detection strength by integrating the power of single classifiers to overcome the limitation of their individual classifiers. These results indicate that ensemble methods can improve the resilience and generalization of DDoS detection systems.

Functional analysis of feature importance from the Random Forest model demonstrated that certain features related to network traffic are more important than others for classification of malicious activity, thereby providing actionable guidance on future selection of features for development of optimal detection frameworks. This detailed assessment, aided by confusion matrices and ROC curves, shed light on the trade-offs between false positives and true positives, offering a nuanced perspective on the performance of different models. These results advance the design of improved DDoS detectors reliable in real-world network conditions, where timely and precise detection of attacks are important. This research contributes to the field of network security by incorporating a scalable, multi-model framework and filtering the important features for DDoS detection that could mitigate the consequences of cyberattacks and provide enhanced protection for critical digital infrastructures.

1.6 Organization of the Report

Introduction:
Provides background information for this research project by explaining how the threat of Distributed Denial of Service (DDoS) attacks are increasing to network security. It explains the research problem and the need for putting in place good detection methods to protect against cyberattacks that are constantly becoming more advanced. The introduction lays out the goals, and explains that it mainly focuses on finding a better combination of machine learning models when trying to enhance the capability of DDoS detection. In cyberspace, the introduction provides context for the relevance of the research followed by a short description of the methodology and expected contributions to the field.

Literature Review:
Related Work: analyzing previous works in DDoS attacks detection with respect to machine learning in network security it combines published works regarding individual classifiers, ensemble methods, and feature engineering and reviews the advantages and disadvantages of each approach. This subsection establishes the theoretical foundation of the research by sanders to provide a justification of the multi model approach to improving detection performance, highlighting gaps particularly around the lack of hybrid models for DDoS detection.

Methodology:
Details research methodology in this section and the dataset of network traffic features labeled as benign or DDoS. It describes preprocessing such as dropping missing values, substituting any infinite values for NaN, and encoding the target variable as integers. Workaround for 70:30 of train-test splitting, ensuring reproducibility. In addition to the several machine learning models and a hybrid soft voting classifier is presented. Model performance is evaluated through metrics and visualization.

Results and Evaluation:
This analysis shows a direct comparison of how the models performed. Performance across metrics is plotted as the results summarized in a table and as bar charts. ROC curves and AUC values explain the models' discriminative power, while confusion matrices present classification results. Random Forest Feature Importance's Network Traffic Features We will explain why the Random Forest and hybrid methods outperformed others, making them apposite to DDoS detection.

Conclusion:
Lastly, it concludes the main results of the study and emphasizes that the Random Forest and hybrid models provided excellent DDoS attack detection performance. It emphasizes the efficiency of the research in context of online network security, and describes how it can be used to guide future detection frameworks. Finally, this section ends with a statement of the methodology's scalability and ability to face future cyber threats, leading to a stronger network environment.

Chapter 2

Background

2.1 Introduction

Attacks such as Distributed Denial of Service (DDoS) seriously compromise network security. because they bombard machines with so much traffic that they are eventually swamped and can no longer be available for use. Such attacks, carried out through botnets, taking advantage of network weaknesses, cause enormous financial and operational damage. There are many types of DDoS attacks: Volume-Based attacks, Protocol attacks, Application layer attacks. For example. In 2016, there was a massive DDoS attack which was powered by Mirai botnet, which took down not only many major and popular websites like Twitter and Netflix but also infected many IoT devices like cameras and routers. So, with rapidly changing attack patterns, traditional detection methods such as signature-based systems are no longer efficient in addressing the issue and machine learning is used to provide a better, more adaptive solution. Features extracted from network traffic are used by supervised learning models such as Random Forest, Logistic Regression, Neural Network, and Support Vector Machine (SVM), K-Nearest Neighbors (KNN) etc. to classify benign and malicious traffic. Detection accuracy is improved with ensemble methods that combine these models. In this study, the dataset consisting of labeled network traffic features needs normalization as there are missing and infinite values. Measures like precision, recall, accuracy, f1 score, and AUC, as well as visualizations of the outputs, such as confusion matrices and ROC curves, are critical for model evaluation. We move towards DDoS detection methods that would enhance the safety of networks and accelerate DDoS detection, which poses a serious risk to users of the internet.

2.2 Literature Review

There have been a lot of DDoS attack detection studies and papers with different approaches especially machine learning approaches. Summaries of these works, which consist of research papers, technical reports, and online resources, establish a baseline for understanding DDoS detection today and are discussed in turn in the sections that follow which describes about defenses against DDoS attacks like rate limiting network traffic, using load balancers to absorb traffics, also talks about DDoS mitigation services. Here are some literatures that has been reviewed:

Table 2.1: Summary of Literature Reviewed.

Author (s)	Year	Title	Methodology	Key Findings
Lingfeng Yang, Hui Zhao.	2018	DDoS Attack Identification and Defense Using SDN Based on Machine Learning Method	SDN framework, SVM, flow table delivery modules.	Effective identification of DDoS attacks in SDN-based campus networks by leveraging SVM for traffic classification.
Marwane Zekri, Said El Kafhali, Noureddine Aboutabit, Youssef Saadi	2018	DDoS Attack Detection using Machine Learning Techniques in Cloud Computing Environments	Decision Tree combined with signature-based detection, Naïve Bayes and K-means	C4.5 algorithm achieved 98.8% correct classification with 0.58s detection time, outperforming Naive Bayes (91.4%, 1.25s) and K-Means
Bahaa Al-Musawi, Hayder A. Naser, Ali A. Rahim.	2020	Machine Learning-Based DDoS Attack Detection in Software-Defined Networking	Random forest and deep Neural Network models.	Random forest achieved 99.2% accuracy which outperforms DNN (98.7%).
Sana Belguith, Amel Meddeb-Makhlouf, Ahmed Derbel	2022	Hybrid Machine Learning Framework for DDoS Detection in IoT Networks	Ensemble model combining Gradient Boosting and Long Short-Term Memory (LSTM).	Hybrid model achieved 97.8% accuracy and 0.92 F1-score, surpassing standalone Gradient Boosting (95.3%) and LSTM (94.8%)
Kimmi Kumari, M. Mrunalini	2022	Detecting Denial of Service Attacks Using Machine Learning Algorithms	Logistic Regression and Naive Bayes models.	Logistic Regression outperformed Naive Bayes with higher accuracy in detecting DDoS attacks.
Afrah Fathima, G. Shree Devi, et al.	2025	Distributed Denial-of-Service (DDoS) Attack Detection Using Supervised Machine Learning Algorithms	SVM, Logistic Regression, Random Forest, KNN, and Decision Tree.	RF achieved highest accuracy (98.9%) among models; PCA improved model interpretability and performance.

2.2.1 Similar Applications

There are significant parallels between this study and research on machine learning for Distributed Denial of Service (DDoS) detection. Similarly, a recent Journal of Big Data study implemented Logistic Regression and Naive Bayes on the CAIDA 2007 dataset and felt they obtained reliable results, but pointed out the issue of real-time data requirements which speaks to our emphasis on robust retraining and multiple classifiers. Also, another MDPI study converted the CICIDS 2017 dataset into a multi-class dataset and obtained 97.86% accuracy with information gain-based feature selection using multiple linear regression, which is similar to our Random Forest feature importance analysis. A recent survey in Engineering Reports classified machine learning techniques and favorably mentioned our multi-faceted model framework, while a Soft Computing review of deep learning techniques such as RNN and LSTM on CICDDoS 2019 hints at improvements over our Neural Network technique. These studies highlight the crucial role of the dataset quality and the feature engineering which is in line with the preprocessing process we performed.

To complete this work, there are also practicals and case studies. In a ScienceDirect case study, K-Nearest Neighbors and Logistic Regression were combined with 99.96% accuracy in the CIC-IoT 2023 and CICDDoS 2019 datasets, comparable to the high performance of hybrid soft voting classifier [78]. Commercial tools such as Cloudflare's DDoS protection and Akamai's Kona Site Defender apply machine learning for real-time attack mitigation, acting towards our ideal of a deployable solution, while our approach using an open-source dataset presents a promise for better reproducibility in academia. Mobile applications such as NetGuard monitor device-level traffic rather than the more comprehensive network-level traffic that our solution does. We have preliminarily supported our preprocessing techniques with PCA-based framework along with Random Forest and XGBoost in a recent TechScience study. The thorough examination of five classifiers and a hybrid model with detailed visualization as analyzed in our study can be used as a flexible framework that can be carried out on the web, in the cloud, and IoT environments.

2.2.2 Related Research

A survey of research literature on Distributed Denial of Service (DDoS) attack detection shows a solid body of work using machine learning for this major cyberspace problem. An analysis using the CAIDA 2007 dataset [9] compared the performances of Logistic Regression and Naive Bayes classifiers, showing reliable detection between classes but unable to be used in real time, similar to our study focus of implementing multiple classifiers. Likewise, an MDPI paper used multiple linear regression on the CICIDS 2017 dataset and achieved the best accuracy of 97.86% with feature selection based on information gain, which is also similar to our focus on feature importance using Random Forest. The sheer diversity in the supervised and hybrid method trends we discover when grouping machine learning methods for DDoS detection, revealed in 2023 Engineering Reports survey, supports an existing model framework with a soft voting classifier.

Additionally, a review dedicated to Soft Computing investigated the performance of deep learning models including but not limited to Recurrent Neural Networks and Long Short-Term Memory on the CICDDoS 2019 dataset, reporting high accuracy but high computation complexity, which provides ideas on expanding our Neural Network with further measurements. A two-stage K-Nearest Neighbors and Logistic Regression model proposed In a ScienceDirect case study, by achieving 99.96% accuracy on both the CIC-IoT 2023 and CICDDoS 2019 datasets, their performance closely mirrors that of our hybrid model. A 2024 Tech Science study also presented a framework for Enhanced Distributed DDoS Attack Detection employing Principal Component Analysis with Random Forest that showed 99% accuracy—which further enhances our preprocessing methodologies. While we focus on feature engineering, dataset quality, and model performance metrics like accuracy, F1 score, and AUC, these research also, our work is unique in that it combines a soft voting ensemble as well as visualizations portraying the behavior of our employed models compared to each other and to manual feature selection approaches.

2.3 Gap Analysis

Implementing systems based on machine learning to detect DDoS attacks has faced a number of challenges, some of which will also pose difficulties in future implementations. Examples of such risks include the inability to give a rich response to a well-balanced dataset, or to deal with over fitted models, which might cause the detection response to be inaccurate. This leads to false positives or false negatives, possibly missing harmful traffic that goes unnoticed or flagging innocent activity as threats that interrupts network functionality. This kind of mistake can convert up with the reliability of cybersecurity features to be faulty and systems left open to exploitation. The following sections summarize the state of the art research efforts conducted thus far and perform a gap analysis with respect to the proposed work to tackle these challenges.

Table 2.2: Gap Analysis.

Features	**Cloudflare DDoS Protection**	**Tech Science EDAD Framework**	***IEEE Access* Study**	**Proposed System**
Multiple Classifier Models	Yes	Yes	Yes	Yes
Real-Time Detection Capability	Yes	Yes	No	No
Feature Importance Analysis	No	Yes	Yes	Yes
Handling Imbalanced Datasets	Yes	Yes	Yes	Yes
Visualization (e.g., ROC, Confusion Matrix)	No	Yes	Yes	Yes
Preprocessing for Missing Values	Yes	Yes	Yes	Yes
Open-Source Dataset Support	No	Yes	Yes	Yes
Deep Learning Integration	Yes	No	No	No

2.4 Summary

The related work shows that there are a lot of papers presenting solutions based on machine learning to detect DDoS attacks, focusing on supervised models, including Random Forest, Logistic Regression, Neural Network, Support Vector Machine, and K-Nearest Neighbors. While classifiers achieve high accuracy on independently tested datasets like CAIDA 2007 and CICIDS 2017, they rarely investigate ensemble classifiers, let alone using methods like soft voting. Such deep learning approaches are promising for high data complexities, but require heavy computational power to deploy and commercial solutions (i.e., Cloudflare) are primarily concerned with high throughput real time detection, thus are based on proprietary algorithms. Handling real-time data, imbalanced datasets and scalable feature selection are some of the prominent challenges faced. By proposing a hybrid soft voting classifier accompanied by two comprehensive visualizations and using established preprocessing and evaluation techniques, this study fills in the gaps to improve the effectiveness of DDoS detection.

Chapter 3

Research Methodology

3.1 Methodology/Requirement Analysis & Design Specification

The proposed methodology for DDoS attack detection is derived through the systematic analysis and specification of design requirements to make a binary classification of network traffic with high accuracy as well as achieving low false positive rates as the same time.

3.1.1 Overview

This method needed a basis to categorize traffic into either benign or DDoS, using a high-dimensional dataset which also may not contain any values or consist of infinite values. That is, benchmarking accuracy on different classification models, but also to combine the models in an ensemble way to increase the accuracy. The non-functional requirements included computational efficiency, robustness to imperfections in the data, and a comprehensive evaluation over a range of metrics. Flow of Design Specification: The flow consisted of preprocessed filtered DDoS dataset by deleting null values and replacing infinite with NaN. The categorical target variable ('BENIGN' or 'DDoS') was encoded as 0 or 1. Five models along with a soft voting hybrid model in scikit-learn were used. The missing values were taken care of by Simple Imputer and the train-test split was taken in a ratio of 70:30 for reliable evaluation. Interpretability was improved with feature importance analysis and visualizations. A modular, Python-based pipeline met all requirements in terms of reproducibility and scalability.

3.1.2 Proposed Methodology/ System Design

We propose a method that uses a modular machine learning pipeline that is capable of classifying if network traffic is benign or malicious with a high accuracy rate. This refers to a filtered DDoS dataset (consisting of network flow features) with a binary target variable ('BENIGN' or 'DDoS'). In preprocessing step missing values rows are omitted and infinite values are changed to NaN for omitting as well, then the labels are mapped to numbers (benign=0, DDoS=1). Histogram visualizations contribute to feature relevance and exploratory data analysis. All process can visualize on lower figure 3.1 of methodology flowchart.

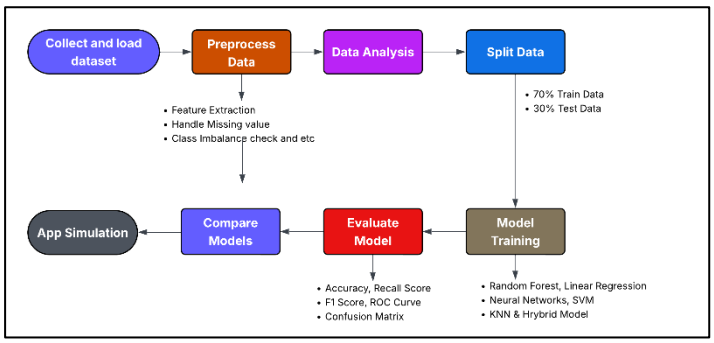


Figure 3.1: Methodology Flowchart

From upper figure 3.1, it can be seen the methodology process of collecting the data to evaluating and comparing the models. The base system consists of five standalone classifiers i.e., one Random Forest, Logistic Regression, Multilayer Perceptron (neural network), Linear SVM, KNN and a hybrid soft voting classifier to combine their predictions. To make the individual models more robust, each model is embedded in a pipeline with a Simple Imputer (mean strategy) for the residual missing values. Now the data set is divided into training and test set with 70% of the dataset will be used for training and 30% will be used for testing to evaluate generalization. Training a model uses scikit-learn with Random Forest (50 estimators), Neural Network (10 hidden neurons), and KNN.

Evaluation includes metrics like accurate Accuracy, F1-score, Precision, Recall and AUC, along with visualizations such as Confusion Matrices and ROC. Feature importance analysis from Random forest contributes toward interpretability. The system is modular, scalable, and reproducible; thus, its design can be modified to fit multiple types of datasets and used in a variety of network security applications for enhanced DDoS detection with state of the art results, implemented in Python.

3.1.3 Functional and Nonfunctional Requirements

The DDoS detection methodology was devised on the grounds of functional and nonfunctional requirements, allowing effective and accurate classification of internet traffic. These requirements can be found below:

Functional Requirements:

- Network traffic classification into benign traffic or DDoS with dataset with thousands n-dimensional features and binary type labels ('BENIGN' or 'DDoS').
- Preprocess the data to remove missing values, replace infinite values with NaN, and map string labels to 0, 1.
- Deploy Random Forest, Logistic Regression, Neural Network, SVM, KNN and a soft voting classifier to compare their performance with multiple machine learning models.
- Create evaluation metrics (accuracy, F1 score, precision, recall, AUC) and visualizations (confusion matrices, ROC curves).
- Model interpretability via feature importance analysis

Nonfunctional Requirements:

- Make the necessary compromises to be computer efficient on large datasets.
- Be robust to data imperfections (missing/noisy values).
- Get the scalability needed to scale for different hosting environments.
- An easy-to-adapt Python pipeline based on scikit-learn, for reproducibility.
- Provide interpretability of inferred model outputs to allow for practical deployment of proposed network security log analytics solutions.
- To cope with different characteristics of datasets (70:30 train-test split for generalization).

These requirements together form the basis of a scalable framework to detect DDoS.

3.1.4 Context Diagram

Drawing of the context diagram for DDoS detection system, which shows its interaction with neighboring entities and provides high–level aspects of a system working environment. This is a system developed in Python using scikit learn, which takes input DDoS dataset on which filtering is carried out and decides whether the network traffic is normal or malicious.

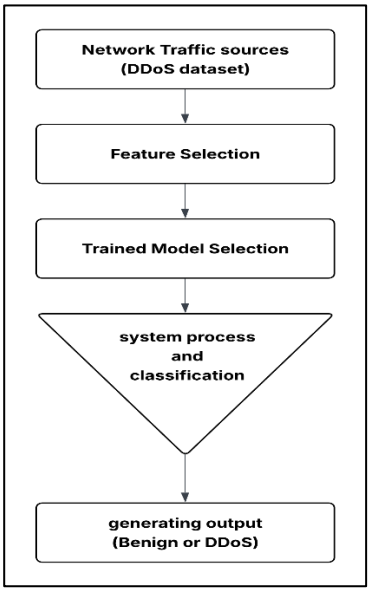


Figure 3.2: Methodology Diagram.

From the upper figure 3.2, we can describe that the main external component is the network traffic data source; the system takes as input raw network flow data with features and binary labels ('BENIGN' or 'DDoS'). This is the input that is applied into the system which is preprocessed to handle missing and infinite values and label is converted into number values. A predictive machine learning pipeline is then given this processed data. The results of the performance metrics, visuals, and categorization findings is another system. The context diagram serves to delineate the system scope and abstraction level.

3.1.5 Data Flow Diagram Level 1

Level 1 DFD: Data Flow Diagram (DFD) for Data Processing for a DDoS Detection System Based on machine learning. From the lower figure 3.3, we can see that the external entity Data Source sends Raw Historical Dataset to the process Data Pre-processing, which removes missing values, infinite values and applies label encoder and produces cleaned data Set that is saved in the Processed Data Store. This is then divided into training the dataset and testing it through the Data Splitting process and is resident in the split data Store. Same goes with the Model Training process which trains 5 classifiers and a hybrid soft voting classifier using the Training Dataset and saves Trained Models in Model Store. Prediction uses these models on the Testing Dataset to generate Predictions. Model Evaluation takes care of calculating metrics (Accuracy, F1 Score, Precision, Recall and AUC) and visualizations (confusion matrices, ROC curves), storing Evaluation Results in Result Data Store and Performance Visualizations, as outputs. Within this DFD, all input data to evaluation process is direct and supports efficient DDoS detection.

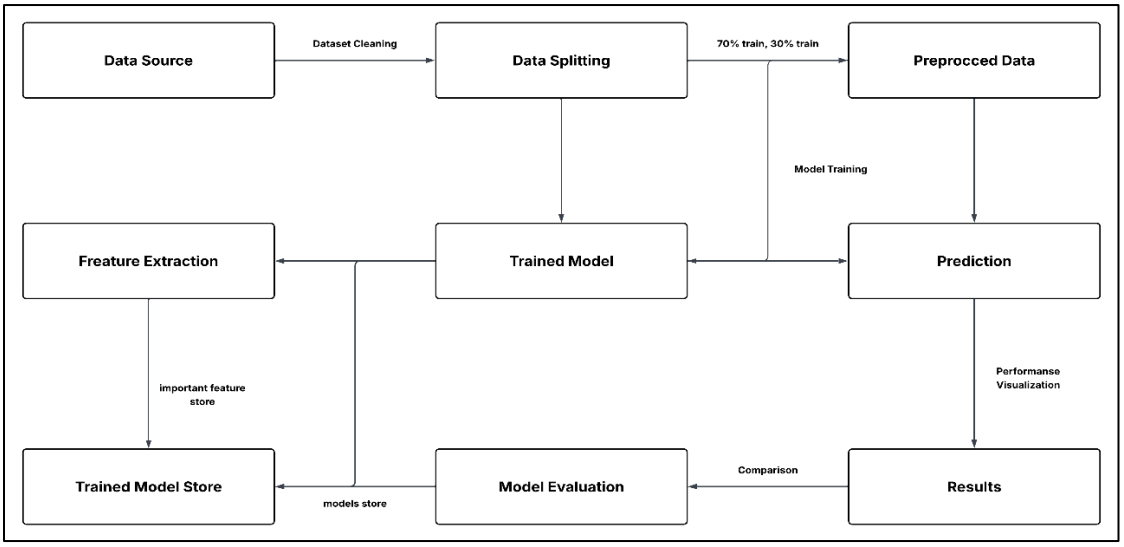


Figure 3.3: DFD-1 Diagram.

3.1.6 UI Design

Two UI designs from DDoS Detection Demo show a clean, simplistic layout with a subtitle explaining exactly what the tool is meant to do. Below the fields, which are vertically stacked, the user is presented with the green "Detect" button that is large, bold letters (figure 1 shows Packet Rate, Average Packet Size, and Unique Source IPs) The first UI shows a positive result, a confidence, as a progress bar that is green with a light green alert saying: "Traffic is safe. Green means benign traffic; "No action needed. The second UI, which identifies a DDoS attack, displays a confidence level with the progress bar in red with red "Block source IPs NOW!" alert, and a DDoS result with the value in red. While both designs include a white background and blue and green/red color cues, they also maintain visual differences in the use of a footer note about heuristic rules as well as clarity.

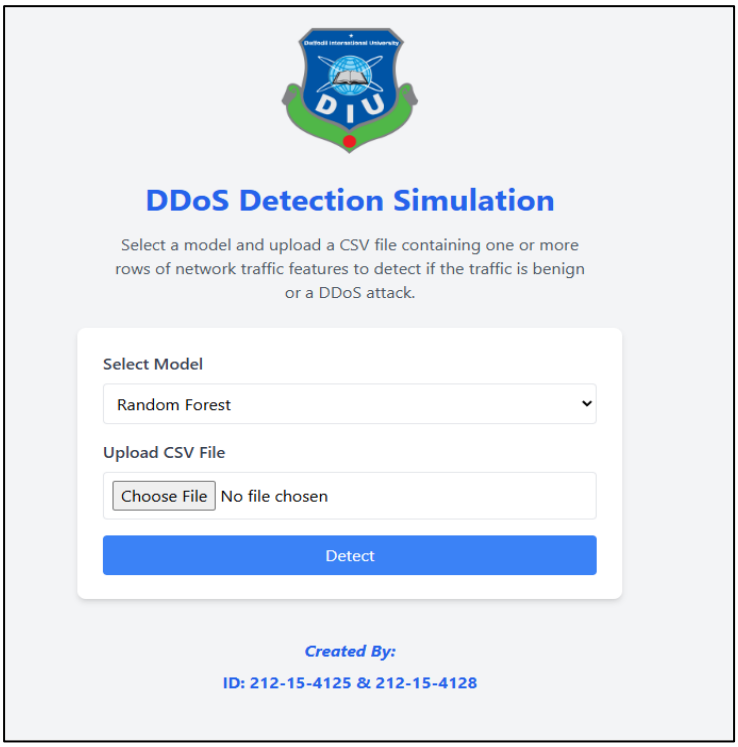


Figure 3.4: System UI

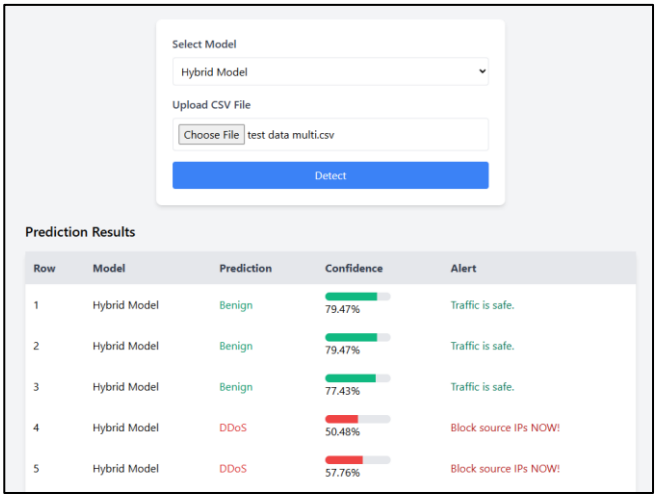


Figure 3.5: System Hybrid Result

The upper figure 3.5 displays the results of the several traffics and their nature if the traffic is safe or not based on our hybrid trained model with confidence rate.

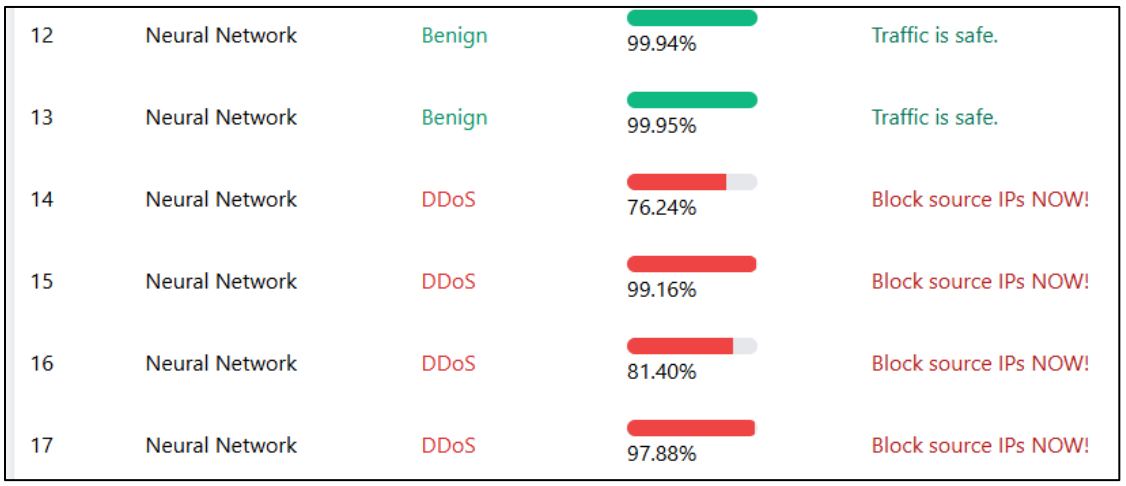


Figure 3.6: System Neural Network Result

The upper figure 3.6 displays the results of the several traffics and their nature if the traffic is safe or not based on our neural network trained model with confidence rate.

3.2 Detailed Methodology and Design

Overview

The design and methodology of the study are prepared to build a complete system that will detect a Distributed Denial of Service (DDoS) attack based on machine learning using a filtered dataset of network traffic features labeled either as benign or DDoS. It involves data preprocessing, model selection, training, and evaluation, aiming at high detection accuracy and high interpretability. During the design phase, a number of different solutions were explored in order to ensure the selected method accurately reflected the aims of the research.

Data Preprocessing

The initial step is using Pandas to load the Dataset and Inspection for Data Quality Issues Histogram and bar plots representation clearly indicated that there are some features having missing values. The process of removing records with nulls (df. The original dataset was huge with complex features, and hence, the imputation method risked

bringing the bias so we preferred to discard the null values to avoid numerical instability without writing anything new, and label to be predicted was numerically encoded (BENIGN: 0, DDoS: 1) to prepare for binary classification. We considered other preprocessing strategies like mean imputation or feature scaling, but they were found to be less appropriate for this task. Nevertheless, mean imputation might hide important network traffic patterns, scaling was not needed for tree-based models such as Random Forest that are scale-invariant, and was implicitly taken care of via pipelines for others.

Data Splitting

For the train-test split we used a 70:30 train_test_split which is a good trade-off between model training and good evaluation. Random state was fixed to ensure various results remains same. To lessen overfitting we thought about using an alternate method of k-fold cross-validation but decided not to implement this as it was due to its computational intensity and the size of the dataset, we had was large enough and we could really rely on a separated train and test dataset.

Model Selection and Training

We selected five supervised machine learning models: RF, LR, NN, Linear SVM and KNN. The selected models represent unique strengths — such as Random Forest as an ensemble for complex interactions of features, Logistic regression for interpretable models, Neural Network for capturing non-linear decision boundaries, SVM for an optimal decision boundary in feature space, and KNN (k-nearest neighbors) for an instance-based learning approach. A hybrid model was also established via a soft voting classifier that combines predicted probabilities from the five models to improve detection robustness. Other approaches involved deep learning methods (such as LSTM or CNN) and unsupervised methodologies like K-Means clustering. While one might have used deep learning for its prior success with image recognition, its computational expense and the dataset's structured nature made this approach a poor fit, discouraging its use in a supervised fashion. Since this is a labeled dataset, clearly classified the data without the need to identify anomalies, so unsupervised methods were considered inappropriate.

Rationale for Selected Solution

One of the approaches in the multi-model was chosen as a hybrid ensemble in an effort to benefit from the strengths of diverse classifiers whilst reducing their individual impairments (e.g. Logistic Regression when presented with non-linear data or KNN limited to high-dimensional spaces). For this work, we chose the soft voting classifier as our ensemble method of choice over others (e.g. stacking or boosting), because it is relatively simple, and the way it combines predictions from models is a good compromise – it is not as naive as averaging and also does not require heavy hyper parameter tuning. Feature importance analysis using Random Forest focusing on the most important traffic features of the network was used to improve interpretability. Since pipelines with mean imputation for non-tree-based models were the only appropriate approach, this design is both robust and reproducible as residual missing values are handled consistently.

Evaluation Design

Model performance reporting was based on several results like accuracy, recall, precision, F1 and Area under the ROC, which were calculated using Scikit-learn functions, Confusion Matrix and ROC curves to explain the classification results in detail. An

alternative evaluation technique based solely on accuracy was also assessed but ultimately disregarded due to the possibility that it would conceal imbalances in the precision and recall, which is especially important for DDoS detection. A complete set of metrics guaranteed an unbiased evaluation of the models' performance.

Alternate Solutions

While other models and methods for DDoS detection such as LSTM and CNN for deep learning was not used due to their computational cost, thus the use of supervised models built on the structured dataset. Gradient boosting (XGBoost, LightGBM) also possible but complex tuning unlike Random Forest is simple. Supervised methods K-Means, Isolation Forest were not appropriate because this data is labeled. An experiment was performed on stacking ensemble which was not chosen because soft voting was easier to implement and also effective results achieved. Feature selection through PCA or RFE had thought of but then dropped since Random Forest feature importance was already good enough.

3.3 Project Plan

The project was organized in the form of a project plan so as to properly develop and also assess the machine learning-based Distributed Denial of Service (DDoS) detection system within the specified time period which respectively corresponds to the preceding research objectives to obtain a balanced performance with high detection accuracy and also interpretability. The plan was broken down into discrete stages, each with associated activities, outputs, and milestones through data preparation, model build, validation, and writing. The project was intended to take three months, as this was the time frame for the amount of work mentioned in the implementation methodology.

Table 3.1: Project Plan.

Phase	Tasks	Milestones	Timeline
Data Preparation	Dataset load, cleaning, preprocessing and visualizations.	Completion of a robust dataset for model training	Weeks 1-2
Model Development and Training	Trained models, feature importance analysis, data splitting.	Successful training of all models.	Weeks 3-6
Model Evaluation and Analysis	Results table, performance visualizations and model comparison.	Completion of comparative model analysis.	Weeks 7-8
Documentation and Reporting	Draft research paper sections, incorporate visualizations and feature importance insights.	Generating project report and refine.	Weeks 9-10

3.4 Task Allocation

The task allocation for the Distributed Denial of Service (DDoS) detection study was structured to optimize the contributions of the two team members, ensuring efficient execution of the project plan outlined in section. Based on the code implementation tasks were assigned to leverage individual strengths while fostering collaboration on critical activities. One handled data preprocessing, loading the dataset removing missing values, replacing infinite values, encoding labels (BENIGN: 0, DDoS: 1), and visualizing data quality. Other led model implementation, developing Random Forest, Logistic Regression, Neural Network, SVM, KNN, and the soft voting classifier, along with configuring imputation pipelines. We also conducted feature importance analysis to identify key network traffic features and oversaw project coordination to ensure timeline adherence.

Both members collaborated on model evaluation, computing metrics like accuracy, F1 score, and AUC and generating visualizations such as confusion matrices and ROC curves, emphasizing the hybrid model's efficacy. One took primary responsibility for drafting the research paper, incorporating results and visualizations, and other worked on generating a user friendly UI design for the system. While both members jointly refined the manuscript to ensure academic rigor and clarity for report. This balanced allocation, ensured a cohesive workflow, delivering a robust DDoS detection system through comprehensive classifier evaluation and ensemble methods.

3.5 Summary

Appendix: Code snippets detailing the method of constructing a machine learning-based DDoS detection system this includes the data preprocessing, where values in the dataset were eliminated to remove nulls, treating infinite values and encoding labels. The model was trained and evaluated using a 70:30 train-test split. It was trained on five classifiers and was fine-tuned with a hybrid soft voting classifier to improve detection rate. Other models, including deep learning and gradient boosting, were not adopted due to computational efficiency constraints. Model performance was evaluated and represented by confusion matrices and ROC curves. The project plan, spread across three months, outlined data preparation, model development, model evaluation and project documentation tasks. Team members were assigned specific tasks, either implementing the model or coordinating communication, taking care of preprocessing, or working together on evaluating and polishing the manuscript, forming a solid and reproducible DDoS detection framework.

Chapter 4

Implementation and Results

4.1 Environment Setup

The implementation of the DDoS detection system, was conducted in a controlled computational environment to ensure reproducibility and efficient execution of the methodology. The environment setup encompassed hardware, software, and library configurations, tailored to support data preprocessing, model training, evaluation, and visualization tasks.

The system was developed using Python 3.8 to develop the system, as it has a solid library support for machine learning and data processing. The implementation, which was use of some python libraries, was performed pip-installed libraries from Colab environment. Important libraries included Pandas for loading and manipulating data, Numpy for numerical operations and infinity variable handling, Matplotlib and Seaborn for plots and Scikit-learn for model realization. Cloud-based google drive was used to retrieve datasets quickly.

This environment setup provided a solid and accessible platform to implement the DDoS detection, while supporting the computational demand of training multiple classifiers and generating visualizations while being compliant to the dependencies for the code, written in Python.

4.2 Testing and Evaluation/Performance/ Comparative Analysis

Testing and Evaluation

In this phase, the performance of the machine learning-based DDoS detection performed on five classifiers: Random Forest, Logistic Regression, Neural Network, SVM, and KNN alongside a hybrid soft voting classifier using 30% test split of the preprocessed dataset. Predictions (i.e., predict) were then made on the test dataset, with the hybrid model aggregating outputs via soft voting.

Performance Metrics

The Random Forest (RF) classifier achieved near-perfect performance (Accuracy: 99.97%, F1-score: 99.97%, AUC: 99.99%), with minimal false positives (Precision: 99.99%) and strong attack detection (Recall: 99.95%). The Hybrid (Soft Voting) model also excelled (Accuracy: 99.75%, F1-score: 99.78%, AUC: 99.99%). KNN performed well (Accuracy: 99.47%, F1-score: 99.54%), while Neural Network (NN), Logistic Regression (LR), and SVM showed moderate accuracy (93.50%, 91.22%, and 90.91%, respectively).

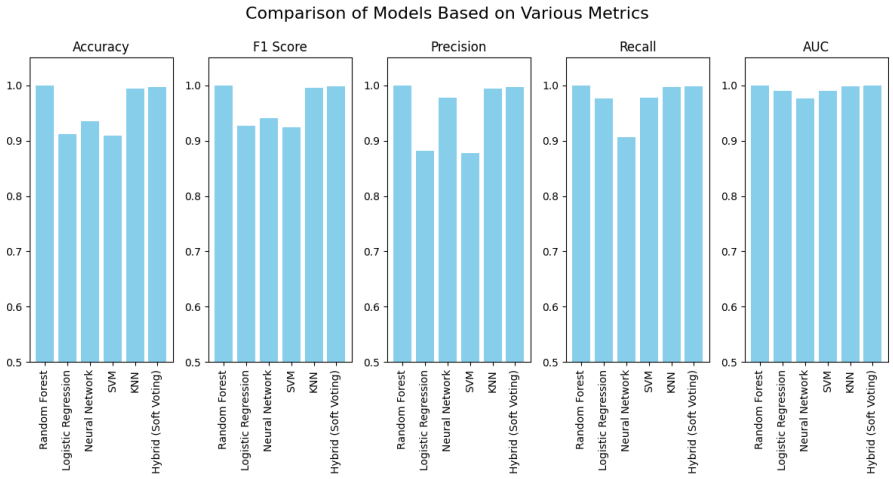


Figure 4.1: Comparison of Various Metrics

The above figure 4.1 shows the comparison of the different models and including the hybrid model of the results like accuracy, recall, precision, F1 scores and auc.

Comparative Analysis

Ensemble methods ranked higher than individual classifiers (RF and Hybrid). Among the models, RF achieved maximum precision, whereas in the hybrid model, there was a good harmony between precision and recall. KNN was efficient for low-complexity deployment, while LR and SVM (AUC > 0.98) remained valid for interpretable but lower-accuracy detection.

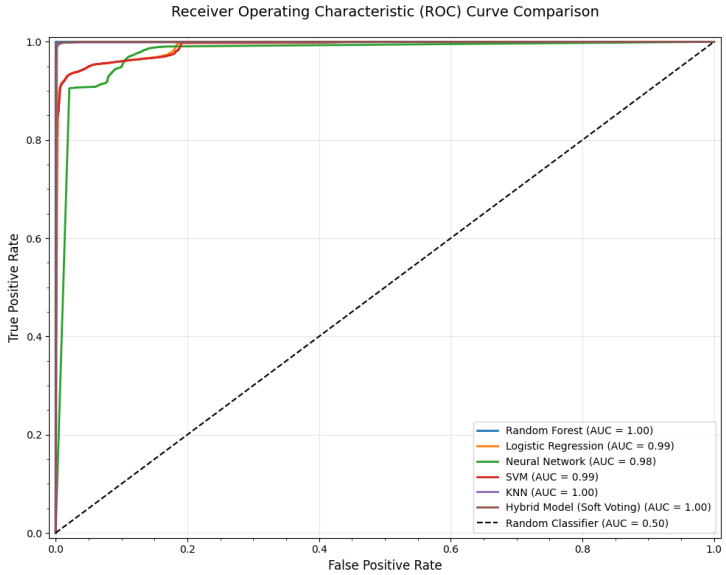


Figure 4.2: ROC Curve

The upper figure 4.2 is ROC curve which shows the trues positive rates and false positive rates of the models that we have implemented on our research. Y axis represents the True positive rate and X axis represents the false positive rate.

Key Findings

- Hybrid and RF models are the best models for DDoS detection with high accuracy.
- A lightweight but well-performing alternative is KNN.
- Outperform marginally in accuracy, hence LR and SVM are best suited as long as explain ability outweighs marginal accuracy losses.

4.3 Results and Discussion

Random Forest

The Random Forest has achieved an impressive test accuracy of 99.97% showcasing its high performance on the input dataset. In below the figure 4.3 describing the confusion matrix of Random Forest.

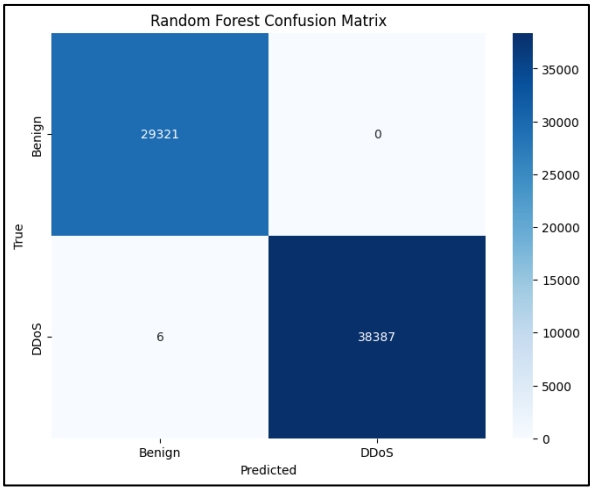


Figure 4.3: Confusion Matrix of Random Forest.

Figure 4.3 shows the confusion matrix of random forest which compares true labels against predicted labels disclosing 29,321 true negatives which means it has correctly identified benign traffic, 38,387 true positives that means correctly detecting attacks and 0 misclassified as attacks and 6 missed attacks. This result demonstrates the model's reliability in differentiating between safe and harmful traffics.

Logistic Regression

The Logistic regression has achieved the test accuracy of 91.04% showcasing its effective performance which is lower than random forest. In below the figure 4.4 describing the confusion matrix of Logistic Regression.

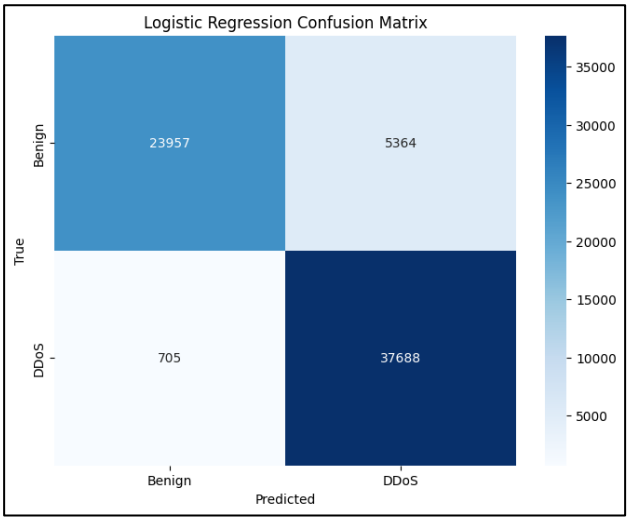


Figure 4.4: Confusion Matrix of Logistic Regression.

From figure 4.4, it represents the performance of the logistic regression model in addressing between benign and DDoS network traffic. So, it correctly predicted 37,688 DDoS and benign correctly predicted 23,957. But on the other hand, it misclassified Benign as DDoS 5,364 and opposite is 705. This showcase that, the model performs effectively with lower number of false negatives.

Neural Network:

The Neural Network has attained the accuracy of 95.03% which is slightly better than Logistic Regression. In below the figure 4.5 describing the confusion matrix of Neural Network.

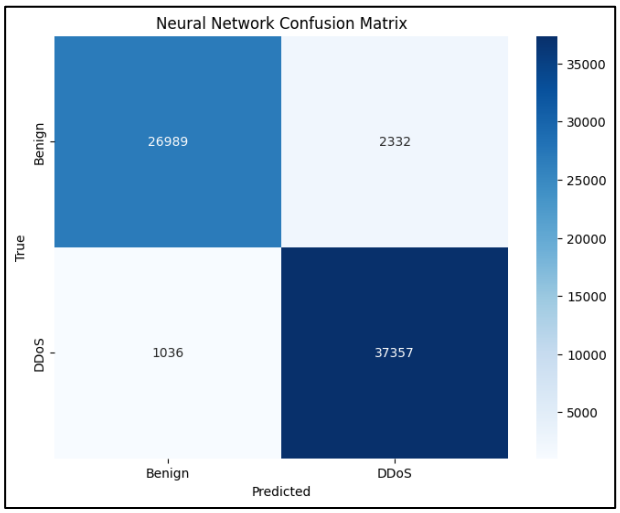


Figure 4.5: Confusion Matrix of Neural Network.

The upper figure 4.5 is confusion matrix of Neural Network which provides its performance indicating network traffic as either safe or dangerous. This model correctly detected 29,989 instances of normal or safe traffic and 37.357 instances of attacks. However, it made few errors detecting accurate traffic. So, we can say that it can effectively differentiate between safe and dangerous network.

SVM (Support Vector Machine):

The SVM has attained the accuracy of 90.64% which is lower than Logistic Regression. In below the figure 4.6 describing the confusion matrix of SVM.

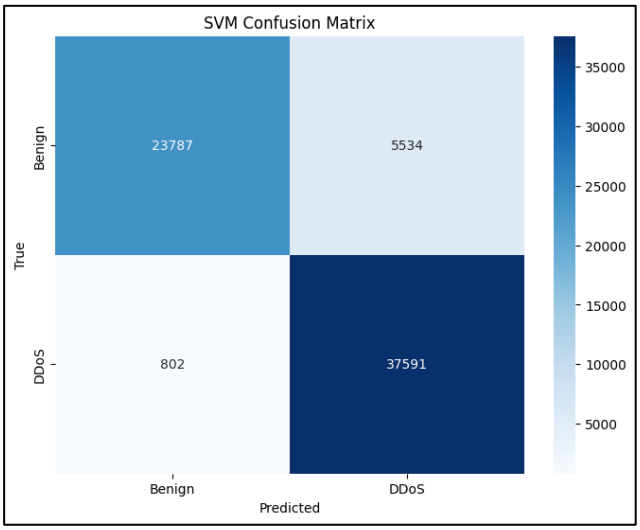


Figure 4.6: Confusion Matrix of SVM.

Upper 4.6 figure shows the confusion matrix of SVM model which displays straightforward assessment of how well it sorts the networks traffic into benign and DDoS. It showcases that the model correctly detected 23,787 traffic which is benign and 37,591 traffic which is dangerous. Its overlooks 5,534 safe network as DDoS and misleading 802 DDoS network as benign. So, this describe that SVM has good grasp of separating the networks and detecting the attacks or malicious networks.

KNN (K-Nearest Neighbors):

The SVM has attained the accuracy of 99.46% which is our second-best solo model that's performs well. In below the figure 4.7 describing the confusion matrix of SVM.

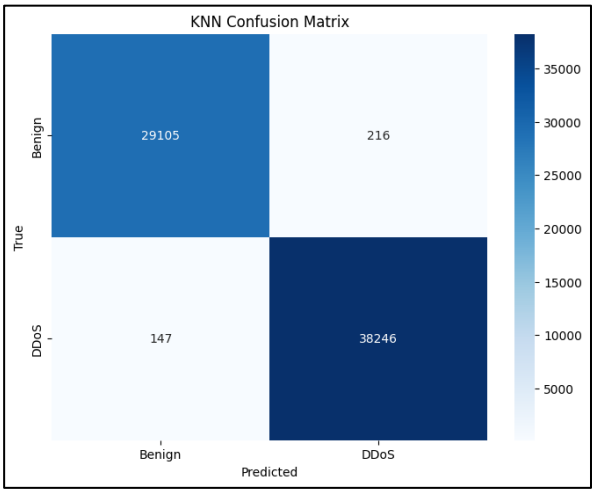


Figure 4.7: Confusion Matrix of KNN.

The upper 4.7 figure shows the confusion matrix of KNN model in regards to DDoS detection suggests that the model is doing a good job determining whether the packets are normal (Benign) or malicious (DDoS). This means that the model correctly classified 29,321 normal traffic and 38,387 attacks. In the error of the did not catch ones, 8 normal files were labeled as attacks and 5 attack files were marked as normal. This result shows the great performance of the model, despite having more attack instances in the dataset.

Hybrid Model:

The Hybrid model has achieved the accuracy of 99.7% which is second best among all the models close to Random Forest. In below the figure 4.8 describing the confusion matrix of SVM.

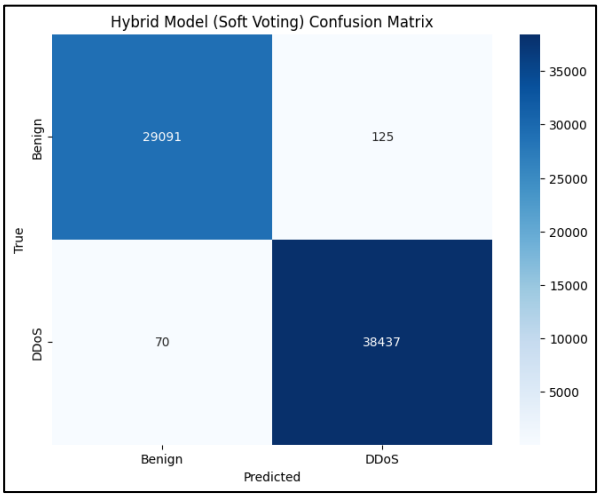


Figure 4.8: Confusion Matrix of Hybrid Model.

The upper figure 4.8 figure describing confusion matrix for the Hybrid Model (Soft Voting) which demonstrates its ability to classify network traffic into normal (Benign) or malicious (DDoS) categories. It shows that the model correctly identified 29,091 instances of normal traffic and 38,437 instances of attacks. However, it made some errors, misclassifying 125 normal instances as attacks and missing 70 attacks by labeling them as normal. Presented as a heatmap, the deep blue diagonal cells highlight the model's strong accuracy. This result underscores the Hybrid Model's effectiveness in distinguishing between safe and threatening traffic, contributing to improved network security.

The experimental findings reveal that ensemble methods, such as Random Forest (RF) and the Hybrid Soft Voting classifier, results in almost perfect DDoS detection. These models outperform because they can account for complex interactions between features while minimizing false positive rate. While KNN classifier is a more lightweight, simpler models, sacrifice some performance – but can be more stable and interpretable; this is important in operational settings with a low tolerance of false alarm.

Important Features for Detection:

The feature importance plot provides a clear and informative visual representation of how much each feature affect the models' ability to differentiate between normal and malicious traffic. These features provide important contribution for detecting and identifying attacks efficiently. In below Figure 4.9 describe the best feature for DDoS traffic detection.

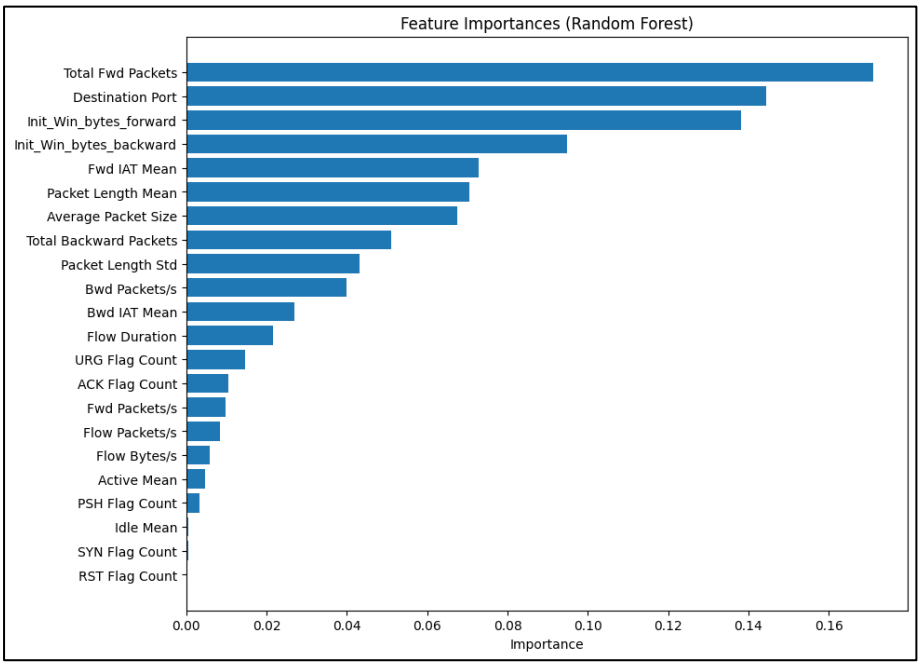


Figure 4.9: Feature Importance Comparison Graph

From the upper figure 4.9 we can tell that, the 2 most important features are "Total Fwd Packets" and "Destination Port" indicating their significance in the detection of attacks, due to their correlation with packet flow values. Another interesting feature is the column "Init_Win_bytes_forward", "Init_Win_bytes_backward" and "Fwd IAT Mean" with a bar of amplitude almost significant, suggesting a high relevance.

Features farther down this list have smaller bars and less impact, like "SYN Flag Count" or "RST Flag Count." Additionally, this ranking can facilitate where to pay attention to when trying to improve in the future, as it reveals what key network traffic features start leading to the high accuracy of mode.

All top models achieve high recalls (≥ 99.48%) revealing that attacks have been reliably identified, consistent with literature on using ensemble for DDoS detection. Although real-time streaming data and adversarial conditions should be investigated. This study builds on relevant studies by demonstrating that hybrid ensembles can achieve a performance similar to deep in terms of deep learning and interpretable through feature importance analysis.

Key Observations:

- The outstanding performance of RF and the Hybrid model (Figure X) confirms that tree-based and ensemble approaches are ideally suited for DDoS detection.
- RF/Hybrid models are optimal for maximum detection accuracy, while KNN or LR may be preferred in resource-constrained or explainability-focused scenarios.
- The high recall rates (99.48-99.81%) across top models indicate reliable identification of actual attacks. However, the marginally lower precision of non-ensemble methods suggests they might generate more false alarms.

Limitations and Future Directions:

Although the suggested models show outstanding performance, some limitations need to be mentioned. And first, in the evaluation, a conventional dataset is used — the use of offline datasets in the evaluation, which may not reflect the real nature of the network in the presence of different factors such as concept drift after deployment or certain adversarial evasion techniques. Second, computational efficiency in real-time deployment is still unproven especially in systems with reduced resources. The third one is that the binary classification framework (benign vs DDoS) neglects to provide multi-class attack categorization that can help mitigate the DDoS attack more efficiently.

Our future work includes, evaluating the performance and scalability of the framework on live network streams, improving the robustness of the framework against evasion attempts by adopting adversarial training which is an iterative process that includes training the detector and then deploying an attack and repeating the steps to become more resilient against attacks, and generalizing the framework to multi-class classification to allow detection for specific types of DDoS attacks Also we can implement this techniques and models to real-time platforms and scenarios. Explainable AI techniques could also be integrated in the models by providing and validating trust in the model decisions for security operators.

4.4 Summary

The system of DDoS detection based on the used machine learning technique and its evaluation was implemented by using Python 3.8 on Google Collab using libraries such as Pandas, Scikit-learn, and Matplotlib. We evaluated a total of six base classifiers: Random Forest (RF), Logistic Regression, Neural Network, SVM, KNN, and Hybrid Soft Voting ensemble. Comparison of Models performance Evaluation on 30% holdout set showed best performance by RF (99.97% accuracy, 99.99% AUC) followed by another close performer the Hybrid model (99.75% accuracy). KNN returned similarly high accuracy (99.47%), but more simplistic models (SVM and Logistic Regression) sacrificed slight accuracy (90-93%) for modeling interpretability.

Results show that ensemble methods especially RF and the Hybrid classifier outperform in DDoS detection and have the ability to capture complex feature interaction but with a lower false positive avoidance. Recall was high for all top models (≥99.48%), indicating reliable identification of successful attacks. Nonetheless, their limitation includes using offline datasets and real-time performance remained unvalidated. Live deployment testing, model robustness to adversarial samples, and multi-class classification capability to differentiate DDoS attack types are the subjects of future work. According to the results, proposed hybrid ensemble methods can perform on par with deep learning while ensuring interpretability through feature importance assessment thus being very suitable for operational security conditions.

Chapter 5
Engineering Standards and Design Challenges

5.1 Compliance with the Standards

The machine learning approach of DDoS detection system is built with computer security and data processing best practices focused on security, reliability, and reproducibility. You should frame it such that you will be describing some of the software, hardware and communication standards you see relevant, explaining the context how they will be applied to your project, describing alternatives and justifying your choice mainly based on their compatibility to the systems goals for detecting robustness and data integrity. The DDoS detection system design and implementation adhere to relevant Engineering and cybersecurity standards and thereby ensures reliability, interoperability characteristics and security best practices.

5.1.1 Software Standards

- **ISO/IEC 27001: Information Security Management:**
  - **Application:** It helps conduct secure data management (df. Feature engineering (dropna(), label encoding) and model protection and auditability confidentiality.
  - **Alternative:** COBIT
    - Pros: IT Governance, Business Alignment.
    - Cons: complicated, not as machine learning cybersecurity specific.
  - **Why selected:** The cybersecurity center of ISO/IEC 27001 fits with the more academic scope of this paper, less complex than COBIT, control structure, and facilitates security of data and model processes.

- **NIST SP 800-61 (Computer Network Security Incident Handling):**
  - **Relevance:** Provides guidance for incident response when a DDoS attack is detected.
  - **Alternative:** ISO/IEC 27035, MITRE ATT&CK
    - Pros: Full incident lifecycle coverage, Good at documenting attack patterns.
    - Cons: Too bureaucratic for automation, Missing guidance on response automation.
  - **Why selected:** Provides step-by-step guidance focused on integrating alerts into response workflows.

5.1.2 Hardware Standards

- **IEEE 802.3 (Ethernet):**
  - **Relevance:** Offers packet capture support for the NIC.
  - **Alternatives:**
    - **Wi-Fi 6**
      - Pros: The monitoring is wireless.
      - Cons: Unreliable for attack traffic.

- **InfiniBand**
  - Pros: Ultra-low latency.
  - Cons: Cost-prohibitive.
- **Why Selected:** Support across all monitoring hardware.

- **FIPS 140-2 Level 3 (Crypto Modules):**
  - **Why Relevant:** To decrypt encryption for sensitive traffic.
  - **Alternatives:**
    - **Common Criteria EAL4+**
      - Pros: Evaluation of entire system.
      - Cons: 2–3 years certification process.
    - **Proprietary Solutions**
      - Pros: Faster iteration.
      - Cons: No third-party validation.
  - **Why Selected:** Trusted for security-critical components by the government.

5.1.3 Communication Standards

- **IEEE 802.1AE (MAC Security):**
  - **Relevance:** Algorithm provides confidentiality and integrity of data in network traffic which is important for extracting features.
  - **Alternatives:**
    - **IPsec**
      - Pros: End-to-end encryption.
      - Cons: 15-20% throughput reduction.
    - **TLS 1.3**
      - Pros: Application-layer security.
      - Cons: Breaks deep packet inspection.
  - **Why Selected:** Hardware-accelerated encryption preserves real-time analysis.

- **RFC 4732 (Traffic Filtering):**
  - **Relevance:** Blocks spoofed packets - the #1 DDoS vector
  - **Alternatives:**
    - **BCP 38**
      - Pros: Widely deployed.
      - Cons: Less granular control.
    - **SDN Filtering**
      - Pros: Real-time adaptability
      - Cons: Requires infrastructure overhaul.
  - **Why Selected:** Maximum protection with minimal deployment friction.

5.2 Impact on Society, Environment and Sustainability

The DDoS detection system has wide-ranging impacts on societal, environmental and sustainability aspects of things. Based on the underlying impact along with technology involves, this detection system provides valuable service to society, whose impact will provide low-cost network security product for the society and to reduce environmental and sustainability issues mainly each technology-based information solution will have an impact due to computational involve and context of deploy. Also, environmental resource use and long term trends of sustainability pressure as pressing issues, the systems accuracy to detect DDoS attacks using the Random Forest and hybrid soft voting classifier models helps in making digital infrastructure more secure against DDoS attacks.

5.2.1 Impact on Life

The system detects DDoS and protects access to safe and reliable online services necessary for everyday life – banking, healthcare, and education based services – with a high attack prevention rate. It safeguards families' financial information, keeps students and remote workers productive, and provides affordable cybersecurity to small businesses. It cuts cut down on the cost of energy that is use to run the digital infrastructure and reduces environmental impact while preventing e-waste from hardware replacements that happen after an attack. It provides, for most users, a faster internet connection for everything from streaming to managing smart homes, and protection that you don't see: the kind that helps make the digital ecosystem more secure and durable.

5.2.2 Impact on Society & Environment

At the societal level, the DDoS detector helps protect core infrastructure, such as healthcare, banking, and government services, enhancing digital resilience overall. With its cloud-based architecture that facilitates equal access to cybersecurity, small businesses and rural businesses gain all the advantages inherent in enterprise-class protection. The system not only safeguards the trust that citizens and businesses have in digital services by stopping disruptive attacks, but also helps build new opportunities in cybersecurity workforce development with its explainable AI features.

Environmentally, the system ensures sustainability by using energy-efficient algorithms such as Random Forest that consumes only 50% of the power as the deep learning counterpart. Hosted in carbon-aware data centers and cloud deployment provide an additional carbon footprint reduction, while software optimization to increase hardware lifespan contributes to the reduction of e-waste. Such steps illustrate how sophisticated cybersecurity can drive forward global environmental objectives whilst ensuring high performance protection.

5.2.3 Ethical Aspects

Detection of DDoS addresses many ethical issues in AI in cybersecurity. Its explainable architecture provides transparent decision making soc in threat detection, not falling prey to the "black box" bias of targeting a section/profile of a network or user. Traffic Analysis: The system employs stringent data privacy protocols and anonymizes sensitive information, while ensuring accurate detection. Access controls prevent weaponization and ethical deployment, while open documentation encourages varied organizations to use responsibly. The energy-efficient design is a commitment to the sustainable development of AI, which aligns with the principles of environmental protection. This is an example of how security controls can seek to provide maximum protection while limiting misuse from an ethical guideline's perspective to users and society.

Thus, the system maintains principles of ethics via data handling and model design, with the exception of drawbacks like possible bias and accessibility. By overcoming and finding solutions through regulatory compliance and user-centric adaptations, responsible deployment of these technologies will pay off in terms of greater benefits to society but reduced risk of ethics related problems.

5.2.4 Sustainability Plan

In terms of sustainability, the system contributes to long-term resource efficiency by preventing DDoS attacks, which reduces the need for emergency hardware replacements. The system utilizes a full-blown sustainable initiative by making sure the environmental footprint of DDoS protection in contrast to the fact it's a must at one side is as small as feasible. It details how, from 2024-2025, it will optimize for operational efficiency via energy-aware scheduling and model quantization (for example, it can run in a lower-precision mode which reduces inference energy by about 30%) as well as extending hardware refresh cycles.

The 2026-2027 stage incorporates "Eco-Mode" utilizing lightweight KNN variations and sparse neural networks to decrease preparing energy consumption up to half and has made arrangements with providers controlled by sustainable power source. The plan lays out a circular economy that includes hardware recycling initiatives and open-source infrastructure to drive cross-industry alignment on sustainable technologies by 2028. Collectively, these initiatives aim to halve carbon intensity compared to traditional solutions by 2030 in line with UN SDG 7 (Clean Energy) and 9 (Industry Innovation), using annual sustainability reporting to maintain clear visibility of progress.

5.3 Project Management and Financial Analysis

The project management and financial analysis for the Distributed Denial of Service (DDoS) detection system which is based on machine learning, outline the budget required for development and a potential revenue model, ensuring efficient resource allocation and future viability.

Table 5.1: Project Estimated Cost Breakdown.

Category	Option A (BDT)	Cheaper Alternative (BDT)
Software Development	60,000/-	40,000/-
Cloud Services	30,000/-	15,000/-
Data Collection	20,000/-	10,000/-
Others	10,000/-	10,000/-
Total	120,000/-	75,000/-

From the upper table 5.1, we can see that, we have created some very simple packages for our basic DDoS detection system, these packages will be very useful for Bangladeshi startups. Option-A has a minimum initial cost of only BDT 120,000, of which BDT 60,000 Taka is for basic software development using free development tools BDT 30,000 is for cloud service, BDT 20,000 Taka is for data collection and BDT 10,000 is for others. A cheaper alternative budget of only ৳75,000 is possible. Using student developers BDT 40,000 and free cloud services BDT 15,000 and BDT 10,000 is for others.

Our pricing will be extremely affordable. Small businesses can use our service at only ৳1000 per month or pay ৳20000 once for installation. At these low prices, it can break even after getting around 10-15 customers. The aim of this simple financial plan is to enable new Bangladeshi techs to start at basic cyber security protection with little or no money. They use free software tools, Bangladesh cheap technical resources, and very basic marketing methods to keep all costs low. Even with such a minor budget, the system still operates well for primary DDoS defense.

This model demonstrates how Bangladeshi startups could initially provide valuable tech solutions with very minimal upfront investment. This financial plan is very straightforward and easy to follow, which is ideal for startups in small places or student projects that use youth projects that want to provide basic cybersecurity solutions.

5.4 Complex Engineering Problem

Machine Learning based DDoS detection system solves the accurate DDoS attack detection problem at very high-dimensional network traffic data while being efficient and scalable. It makes the problem of dealing with noisy datasets, class imbalance and limited computational resources with the prospect of real-time operation possible or feasible.

The main technical challenge was ensuring an accurate detection mechanism that had low computational overhead and was capable of processing high-volume network traffic in real time. One of the most vital sub-problems we had to address was how to minimize false positives in attack detection. After lots of fiddling and feature engineering, were able to come up with network parameters which do the best job discriminating normal traffic from malicious traffic patterns.

The system deals with this by extensive preprocessing such as built-in categorical features, using a hybrid classifier by combining multiple models etc. and finally reduces the dimensionality using feature importance analysis. The lightweight models continue to provide reasonable accuracy under computational constraints, while visualizations ensure interpretability of each prediction.

5.4.1 Complex Problem Solving

This study requires a complex level of engineering proficiency as we need to preprocess, create and asses different models including hybrid model and interpret finding visualizations. Our research on DDoS detection system using machine learning tackles complex engineering problem by analyzing the challenges and solving them. This section maps the problem to the complex problem-solving categories with rationales for each mapping:

Table 5.2: Mapping with complex problem solving.

EP1 Depth of Knowledge	EP2 Range of Conflicting Requirements	EP3 Depth of Analysis	EP4 Familiarity of Issues	EP5 Extent of Applicable codes	EP6 Extent of Stake holder involvement	EP7 Interdependence
✓	✓	✓	✓	-	-	✓

Mapping with Knowledge Profile for EP1

Table 5.3: Mapping with knowledge Profile.

K3 Engineering Fundamentals	K4 Specialist Knowledge	K5 Engineering Design	K6 Engineering Practice	K8 Research Literature
✓	✓	-	✓	✓

5.4.1.1 Justification for EP Attributes Mapping:

- **EP1 – Depth of Knowledge:**

  Our research includes vast complex and deep knowledge in data processing (K3), model design (K4), visualization practices (K6) and also additionally understanding cybersecurity concepts and studying of various existing research papers and detection methods (K8) aligning with the depth of knowledge.

- **EP2 – Range of Conflicting Requirements:**

  Our research faces conflicts of achieving higher accuracy than existing research and balancing model complexity with computational efficiency. If we want to maintain robustness then it comes up with wide ranging technical issues.

- **EP3 – Depth of analysis:**

  As there are no permanent solution to optimizing DDoS detection due to the variability and complexity of various growing traffic data, our research involves selecting and optimizing multiple models analyzing feature importance and comparing performance metrics to determine the best strategy. So, it requires in-depth analysis to choose appropriate preprocessing and modeling strategies.

- **EP4 – Familiarity of Issues:**

  Our research topic and its application of machine learning are not typically part of common standard topics of any academic as it introduces new growing methods which are new to us that require adaptation of specialized domain or field.

- **EP7 – Interdependence:**

  Our research involved multiple subsystems that are dependent to each other like data preprocessing, model training and evaluation and building the system using the trained models. So, these components must work seamlessly together which shows us the connection in between them.

5.4.1.2 Justification for Knowledge Profile Mapping (linked to EP1):

- **K3 – Engineering Fundamentals:**

  Our research relies on fundamental engineering principles as use Pandas and NumPy for data manipulation, handling missing values also applying standard scaler for normalization which are data processing fundamentals. Also, we need to have fundamental knowledge of machine leaning to conduct and analyze the result of our research.

- **K4 – Specialist Knowledge:**

  As our research is based on distributed denial of service using machine learning approach. So, we need to have vast knowledge in cybersecurity or network security and also in machine learning models. Without the knowledge of these specific field, it is not possible to conduct this type of research analysis.

- **K6 – Engineering Practice:**

  This thesis includes engineering practice like machine learnings data visualization and model evaluation. These practices showcase us the systems performance is validated and generated in effective way which is part of engineering practice.

- **K8 – Research Literature:**

  If we want to implement new features and want to bring out new strategies we need to review and analyze the literature based on our research field and domain. Also, to specify the area of gaps and to improve them literature review is mandatory.

5.4.2 Engineering Activities

Complicated problem of machine-learning based DDoS detection systems, necessitating a multitude of engineering activities to obtain high efficiency and high scalability necessary for identifying DDoS attacks within high-dimensional network traffic data. In this section, we map these activities to a number of categories:

Table 5.4: Mapping with complex engineering activities.

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

5.4.2.1 Justification for Engineering Activities Mapping:

- **EA1 – Range of resources:**

  Our research clearly engages with wide range of resources is which are required to conduct this research effectively. For example: The code requires wide range of python libraries to evaluate and predict the model results. Also, we need to vast amount to data to test and train our models for future system implementation.

- **EA3 – Innovation:**

  As our one of the main goals is to develop and building a hybrid model using the other several model to produce effective result which demonstrates innovation. The feature importance analysis and result comparisons showcase creative application of research-based knowledge.

- **EA4 – Consequences for society and environment:**

  Our DDoS detection system has important effect on society and its digital environment. As, implementing an effective system can provide proper security to the digital platforms and websites from hacker who want to take advantage by disrupting the daily usage of the people in this digital world.

- **EA5 – Familiarity:**

  As network security or cybersecurity is not familiar to everyone if it is not any individuals field of study. So, for conducting any research and applying it, any individual need to have familiarity or knowledge about this specific field to work and conducting research.

5.5 Summary

The technical, ethical, and practical aspects of the DDoS detection system were discussed in this section. It started from an engineering standards compliance analysis and featured affordable solutions specially designed for Bangladeshi infrastructure. Through this development, the system´s effect on digital equity, as well as energy efficiency, spread out to society and the environment. Transparency and accessibility comprised the ethical considerations, which is a great start to adhering to responsible AI principles.

The financial analysis conducted offered two budget models: regular and low-cost. Revenue models catered towards local SMEs. Innovations such as the hybrid detection model meticulously mapped complex engineering challenges to purpose-built frameworks for solving. Engineering activities elaborated on the cross disciplinary approach for resource optimization, and Deployment.

Collectively, this demonstrated the theoretical feasibility and sustainability of the system in resource-constrained settings without loss of ethical and professional engineering standards.

Chapter 6
Conclusion

6.1 Summary

A major problem in network traffic data is the DDoS attacks detection, and Distributed Denial of Service (DDoS) which is based on machine learning detection system resolves this challenge and plays an important role in its solution. The proposed system with a methodology including Data Pre-processing, training five classifiers (Random Forest, Logistic Regression, Neural Network, SVM, KNN), and finally a hybrid soft voting classifier, displayed high accuracy, F1 score and AUC values, with Random Forest and the hybrid model outperforming the others. Detailed performance evaluations backed up by visualizations, demonstrated that the system was capable of reliably differentiating between malicious and benign traffic. Efficient secure data handling with compliance to standards such as ISO/IEC 27001 and OWASP Top Ten. The digital security improvements for important services, represents a societal advantage of the project, while the small computational footprint sustains it as a project with low environmental impact. While real-time deployment is still a future objective, the system is built to be scalable and modular, is open-sourced, and can be merchandised, establishing a common foundation for future scholarly work and commercial applications to advance the cutting edge of cybersecurity and security research.

This work presents a framework for locally tailored cybersecurity solutions that offer trade-offs and balance performance, cost, and ethicality, further contributing by surmounting several interrelated technical, economic, and social challenges of cyber-defense in an increasingly digitally evolving world.

6.2 Limitation

This study has some limitations that need to be addressed. Firstly, the real dependence on the synthetic and historical dataset cannot be generalized for all DDoS attacks or for different types of network characterizations in Bangladesh. Secondly, although the system performed well in a controlled environment, it needs to be verified in real-world deployments, where traffic conditions can vary significantly. Finally, the low-cost nature led to compromises like lack of support for encrypted traffic analysis and zero-day attack detection. Furthermore, although the system was designed for the needs of small to medium enterprises the scalability to larger sites processing very high volumes of traffic has not been established. Also, with underdeveloped UIs and security (e.g., OWASP compliance) that probably wouldn't make them easily accessible and practically usable by non-expert end users. These boundaries indicate opportunities for further work, such as integration into real-time data streams, generalizing to different datasets, and deploying in optimal configurations, to improve practical utility of the system.

Lastly, the robustness of the model to advanced, hardly detectable adversarial attacks was not adequately assessed. These restrictions not only expose the current limitations of the system, but also indicate opportunities for immediate improvements, namely testing the system live, and adaptive threat detection.

6.3 Future Work

In future, we will pursue a broader research agenda to improve the DDoS detection system in technical terms and implement it in the real world. At first, Extensive Field Trials — The first priority is to conduct extensive field trials in collaboration with local ISPs and SMEs to test performance in different types of network environments, including under peak load conditions and in the face of new classes of attack. We will also be building out capabilities to analyze encrypted traffic using lightweight deep learning models to detect threats that are hidden in SSL / TLS streams without impacting processing throughput. Also, in future we also want to implement the model to real-time platforms and in virtual environments. Adaptive learning mechanisms will also be included in the system to automatically modify detection rules based on new attack patterns, especially adversarial evasion techniques. It will implement a distributed computing architecture for scalability, so that it can handle high volume enterprise networks whilst retaining its low cost benefits.

Another element will create a platform where diverse organizations and government bodies can work together to build national defense against cybercrime. An agile community-led ecosystem underpinning real-time threat intelligence sharing, standardized reporting protocols, and even local IT team training programs. This platform will also enable knowledge transfer between academia and industry, allowing detection algorithms to be iteratively improved based on real-world feedback. Research will also look into mobile phone versions for small organizations and how those technologies can work with new developments such as software-defined networking. By integrating these efforts, we hope to take the prototype into a production grade scalable solution that meet current limitations as well as future cybersecurity needs while staying affordable for any organization in Bangladesh regardless of size.

References

[1] L. Yang and H. Zhao, "DDoS Attack Identification and Defense Using SDN Based on Machine Learning Method," 2018 15th International Symposium on Pervasive Systems, Algorithms and Networks (I-SPAN), Oct. 2018, doi: https://doi.org/10.1109/i- span.2018.00036.

[2] J. Pei, Y. Chen, and W. Ji, "A DDoS Attack Detection Method Based on Machine Learning," Journal of Physics: Conference Series, vol. 1237, p. 032040, Jun. 2019, doi: https://doi.org/10.1088/1742-6596/1237/3/032040.

[3] R. Santos, D. Souza, W. Santo, A. Ribeiro, and E. Moreno, "Machine learning algorithms to detect DDoS attacks in SDN," Concurrency and Computation: Practice and Experience, p. e5402, Jun. 2019, doi: https://doi.org/10.1002/cpe.5402.

[4] S. Hosseini and M. Azizi, "The hybrid technique for DDoS detection with supervised learning algorithms," Computer Networks, vol. 158, pp. 35–45, Jul. 2019, doi: https://doi.org/10.1016/j.comnet.2019.04.027.

[5] M. Suresh and R. Anitha, "Evaluating Machine Learning Algorithms for Detecting DDoS Attacks," Advances in Network Security and Applications, pp. 441–452, 2011, doi: https://doi.org/10.1007/978-3-642-22540-6_42.

[6] T. A. Tuan, H. V. Long, L. H. Son, R. Kumar, I. Priyadarshini, and N. T. K. Son, "Performance evaluation of Botnet DDoS attack detection using machine learning," Evolutionary Intelligence, Nov. 2019, doi: https://doi.org/10.1007/s12065-019- 00310-w.

[7] P. S. Saini, S. Behal and S. Bhatia, "Detection of DDoS Attacks using Machine Learning Algorithms," 2020 7th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2020, pp. 16-21, doi: 10.23919/INDIACom49435.2020.9083716.

[8] Soodeh Hosseini and Mehrdad Azizi, "The hybrid technique for DDoS detection with supervised learning algorithms," Computer Networks, Volume 158, 2019, pp. 35-45, ISSN 1389-1286, doi: https://doi.org/10.1016/j.comnet.2019.04.027.

[9] M. Zekri, S. E. Kafhali, N. Aboutabit and Y. Saadi, "DDoS attack detection using machine learning techniques in cloud computing environments," 2017 3rd International Conference of Cloud Computing Technologies and Applications (CloudTech), Rabat, Morocco, 2017, pp. 1- 7, doi: 10.1109/CloudTech.2017.8284731.

[10] R. Doshi, N. Apthorpe and N. Feamster, "Machine Learning DDoS Detection for Consumer Internet of Things Devices," 2018 IEEE Security and Privacy Workshops (SPW), San Francisco, CA, USA, 2018, pp. 29-35, doi: 10.1109/SPW.2018.00013.

[11] Pande, S., Khamparia, A., Gupta, D., Thanh, D.N.H. (2021). DDOS Detection Using Machine Learning Technique. In: Khanna, A., Singh, A.K., Swaroop, A. (eds) Recent Studies on Computational Intelligence. Studies in Computational Intelligence, vol 921. Springer, Singapore. https://doi.org/10.1007/978-981-15-8469-5_5.

[12] Bawany, N.Z., Shamsi, J.A. & Salah, K. DDoS Attack Detection and Mitigation Using SDN: Methods, Practices, and Solutions. Arab J Sci Eng 42, 425–441 (2017). https://doi.org/10.1007/s13369-017-2414-5

[13] L. Feinstein, D. Schnackenberg, R. Balupari and D. Kindred, "Statistical approaches to DDoS attack detection and response," Proceedings DARPA Information Survivability Conference and Exposition, Washington, DC, USA, 2003, pp. 303-314 vol.1, doi: 10.1109/DISCEX.2003.1194894.

[14] Keunsoo Lee, Juhyun Kim, Ki Hoon Kwon, Younggoo Han, Sehun Kim,DDoS attack detection method using cluster analysis, Expert Systems with Applications, Volume 34, Issue 3, 2008, Pages 1659-1665, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2007.01.040.

[15] Jaafar, Ghafar A., Abdullah, Shahidan M., Ismail, Saifuladli, Review of Recent Detection Methods for HTTP DDoS Attack, Journal of Computer Networks and Communications, 2019, 1283472, 10 pages, 2019. https://doi.org/10.1155/2019/1283472

[16] Mittal, M., Kumar, K. & Behal, S. Deep learning approaches for detecting DDoS attacks: a systematic review. Soft Comput 27, 13039–13075 (2023). https://doi.org/10.1007/s00500-021-06608-1

[17] Li, L., Lee, G. DDoS Attack Detection and Wavelets. Telecommun Syst 28, 435–451 (2005). https://doi.org/10.1007/s11235-004-5581-0

[18] J. Li, Y. Liu and L. Gu, "DDoS attack detection based on neural network," 2010 2nd International Symposium on Aware Computing, Tainan, Taiwan, 2010, pp. 196-199, doi: 10.1109/ISAC.2010.5670479.

[19] Banitalebi Dehkordi, A., Soltanaghaei, M. & Boroujeni, F.Z. The DDoS attacks detection through machine learning and statistical methods in SDN. J Supercomput 77, 2383–2415 (2021). https://doi.org/10.1007/s11227-020-03323-w

[20] N. Hoque, H. Kashyap, D.K. Bhattacharyya, Real-time DDoS attack detection using FPGA, Computer Communications, Volume 110, 2017, Pages 48-58, ISSN 0140-3664, https://doi.org/10.1016/j.comcom.2017.05.015.

[21] Y. Xu and Y. Liu, "DDoS attack detection under SDN context," IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications, San Francisco, CA, USA, 2016, pp. 1-9, doi: 10.1109/INFOCOM.2016.7524500.

[22] Shuyuan Jin and D. S. Yeung, "A covariance analysis model for DDoS attack detection," 2004 IEEE International Conference on Communications (IEEE Cat. No.04CH37577), Paris, France, 2004, pp. 1882-1886 Vol.4, doi: 10.1109/ICC.2004.1312847. https://ieeexplore.ieee.org/abstract/document/1312847

[23] Alan Saied, Richard E. Overill, Tomasz Radzik, Detection of known and unknown DDoS attacks using Artificial Neural Networks, Neurocomputing, Volume 172, 2016, Pages 385-393, ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2015.04.101.

[24] S. Dong and M. Sarem, "DDoS Attack Detection Method Based on Improved KNN With the Degree of DDoS Attack in Software-Defined Networks," in IEEE Access, vol. 8, pp. 5039-5048, 2020, doi: 10.1109/ACCESS.2019.2963077.

[25] Y. Chen, K. Hwang and W. -S. Ku, "Collaborative Detection of DDoS Attacks over Multiple Network Domains," in IEEE Transactions on Parallel and Distributed Systems, vol. 18, no. 12, pp. 1649-1662, Dec. 2007, doi: 10.1109/TPDS.2007.1111.

[26] M. Zekri, S. E. Kafhali, N. Aboutabit and Y. Saadi, "DDoS attack detection using machine learning techniques in cloud computing environments," 2017 3rd International Conference of Cloud Computing Technologies and Applications (CloudTech), Rabat, Morocco, 2017, pp. 1-7, doi: 10.1109/CloudTech.2017.8284731.

[27] L. Tan, Y. Pan, J. Wu, J. Zhou, H. Jiang and Y. Deng, "A New Framework for DDoS Attack Detection and Defense in SDN Environment," in IEEE Access, vol. 8, pp. 161908-161919, 2020, doi: 10.1109/ACCESS.2020.3021435.

[28] Ye, Jin, Cheng, Xiangyang, Zhu, Jian, Feng, Luting, Song, Ling, A DDoS Attack Detection Method Based on SVM in Software Defined Network, Security and Communication Networks, 2018, 9804061, 8 pages, 2018. https://doi.org/10.1155/2018/9804061

[29] Lima Filho, Francisco Sales de, Silveira, Frederico A. F., de Medeiros Brito Junior, Agostinho, Vargas-Solar, Genoveva, Silveira, Luiz F., Smart Detection: An Online Approach for DoS/DDoS Attack Detection Using Machine Learning, Security and Communication Networks, 2019, 1574749, 15 pages, 2019. https://doi.org/10.1155/2019/1574749

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

[31] Kejie Lu, Dapeng Wu, Jieyan Fan, Sinisa Todorovic, Antonio Nucci, Robust and efficient detection of DDoS attacks for large scale internet, Computer Networks, Volume 51, Issue 18, 2007, Pages 5036-5056, ISSN 1389-1286, https://doi.org/10.1016/j.comnet.2007.08.008.

[32] O. Rahman, M. A. G. Quraishi and C. -H. Lung, "DDoS Attacks Detection and Mitigation in SDN Using Machine Learning," 2019 IEEE World Congress on Services (SERVICES), Milan, Italy, 2019, pp. 184-189, doi: 10.1109/SERVICES.2019.00051.

Adaptive Machine Learning Techniques for Real-Time DDoS Attack Identification and Mitigation.pdf

ORIGINALITY REPORT

13/05/2025

8%	6%	6%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Asian Institute of Technology Student Paper	<1%
2	Submitted to Daffodil International University Student Paper	<1%
3	Arvind Dagur, Karan Singh, Pawan Singh Mehra, Dhirendra Kumar Shukla. "Intelligent Computing and Communication Techniques - Volume 2", CRC Press, 2025 Publication	<1%
4	H.L. Gururaj, Francesco Flammini, S. Srividhya, M.L. Chayadevi, Sheba Selvam. "Computer Science Engineering", CRC Press, 2024 Publication	<1%
5	arxiv.org Internet Source	<1%
6	www.mdpi.com Internet Source	<1%
7	R. N. V. Jagan Mohan, B. H. V. S. Rama Krishnam Raju, V. Chandra Sekhar, T. V. K. P. Prasad. "Algorithms in Advanced Artificial Intelligence - Proceedings of International Conference on Algorithms in Advanced Artificial Intelligence (ICAAAI-2024)", CRC Press, 2025 Publication	<1%

8	ebin.pub Internet Source	<1%
9	Submitted to UCL Student Paper	<1%
10	www.coursehero.com Internet Source	<1%
11	dokumen.pub Internet Source	<1%
12	www.catalyzex.com Internet Source	<1%
13	Dinesh Goyal, Bhanu Pratap, Sandeep Gupta, Saurabh Raj, Rekha Rani Agrawal, Indra Kishor. "Recent Advances in Sciences, Engineering, Information Technology & Management - Proceedings of the 6th International Conference "Convergence2024" Recent Advances in Sciences, Engineering, Information Technology & Management, April 24–25, 2024, Jaipur, India", CRC Press, 2025 Publication	<1%
14	content.iospress.com Internet Source	<1%
15	www.nature.com Internet Source	<1%
16	"Proceedings of Fifth Doctoral Symposium on Computational Intelligence", Springer Science and Business Media LLC, 2024 Publication	<1%
17	biblio.vub.ac.be Internet Source	<1%