

**Computing and Digital Technologies Project**

**LD7083**

**Dissertation Report**

**On**

**Title: “Classification of DNS attack types using machine learning algorithms on the CIC-Bell-DNS-EXF-2021 dataset”**

|  |  |
| --- | --- |
| **Student Name:** |  |
| **Student Id:** |  |
| **Supervisor Name:** |  |
| **Submission Date:** | 05/09/2024 |

# Declarations

I hereby declare that this research thesis titled “*Classification of DNS attack types using machine learning algorithms on the CIC-Bell-DNS-EXF-2021 dataset*” was conducted by my own research and no one contributed to this research. All the tasks and activities throughout this research thesis are performed by me. I have only taken references from the works of other Authors that are relevant and all these sources are appropriately cited. This work not have been previously submitted by anyone in part or full for any professional course or degree at any university or academic institution.

I am aware of the rules, regulations, and penalties of academic integrity. All the provided content is free from plagiarism set forth by my university. All the used sources are properly cited and referenced to acknowledge their contributions to this research.

# Acknowledgments

First, I would like to provide my heartfelt gratefulness to my Supervisor [Name] for his invaluable learning, guidance, mentorship, and support throughout this whole research. His profound knowledge, guidance, and insightful feedback significantly contributed to shaping this research and kept me motivated to continuously push my abilities and boundaries to perform better. I would like to extend this gratitude to my university and staff who either directly or indirectly supported me during this research.

I am extremely obliged to the authors of the CIC-Bell-DNS-EXF-2021 dataset, which provided a solid base to conduct this research. Without this dataset, I would have been not able to conduct this research including the empirical aspects of this research field.

A special thanks go to all my friends and fellow students, who shared necessary resources, moral support, and constructive feedback at different stages of this research.

Finally, I also want to acknowledge the unwavering love, patience, support, and belief in my abilities that was a constant source to keep me motivated and successfully complete this research.

# Abstract

Thousands of websites are there that every day collect use data via login credentials. The availability of a vast number of networks makes this process more complex and the safety and reliability of these websites might be compromised. Here cybersecurity plays a crucial role in protecting users and organizations from cyber attacks. The Domain Name System (DNS) is the most important and ubiquitous protocol to facilitate seamless network communications. It plays a crucial role in improving the security of communication networks and is based on malevolent domain names. These domain names can be used by firewalls to block malicious communication. This research mainly aims to provide a classification of the types of DNS attacks using machine learning algorithms on the CIC-Bell-DNS-EXF-2021 dataset. This research used three main machine learning algorithms, namely Random Forest, Decision Tree, and Naïve Bayes. The performance of these algorithms is calculated using key performance metrics, such as precision, accuracy, recall rate, and F1 score. The obtained research results demonstrated that the Random Forest machine learning model outperformed other algorithms. The Random Forest model achieved an accuracy of 97%, precision of 95.5%, recall rate of 94.5%, and F1-score of 95%. The key features of the Bell-DNS-EXF-2021 dataset are entropy domain name, no. of queries per second, response time, packet length, and TTL values to differentiate the benign and malicious network traffic.

Contents

[Abstract 3](#_Toc176035489)

[1 Introduction 7](#_Toc176035490)

[1.1 Context and Potential Benefits of Research 7](#_Toc176035491)

[1.2 Aim and Objectives 8](#_Toc176035492)

[1.2.1 Aim 8](#_Toc176035493)

[1.2.2 Objectives 8](#_Toc176035494)

[1.3 Changes to the Research Objectives 8](#_Toc176035495)

[1.4 Professional, Legal, Ethical, and Social Issues 8](#_Toc176035496)

[1.5 Report Structure 9](#_Toc176035497)

[1.6 Summary 10](#_Toc176035498)

[2 Literature Review 11](#_Toc176035499)

[2.1 Introduction 11](#_Toc176035500)

[2.2 Evolution of Domain Name System 11](#_Toc176035501)

[2.3 Cyber Threats or Attacks to DNS and Mitigation 12](#_Toc176035502)

[2.4 Cloud-based Intrusion Detection System to Enhance DNS Security 13](#_Toc176035503)

[2.5 Malicious Network Traffic Detection for DNS 14](#_Toc176035504)

[2.6 Malicious Activity Detection in DoH Protocol 14](#_Toc176035505)

[2.7 ML Classifiers to Detect Malicious Activities across DOH Traffic 14](#_Toc176035506)

[2.8 A Unified Approach for Domain Name Detection 15](#_Toc176035507)

[2.9 Semi-Supervised Learning and Parameter Optimization for Malicious Domain Detection 16](#_Toc176035508)

[2.10 Research Gaps 17](#_Toc176035509)

[2.11 Summary 17](#_Toc176035510)

[3 Methodology and ‘Design of Practical Work 18](#_Toc176035511)

[3.1 Introduction 18](#_Toc176035512)

[3.2 Methodology 18](#_Toc176035513)

[3.2.1 Research Approach 19](#_Toc176035514)

[3.2.2 Data Preparation 19](#_Toc176035515)

[3.2.3 Feature Engineering 22](#_Toc176035516)

[3.2.4 Model Selection 22](#_Toc176035517)

[3.2.5 Data Splitting 23](#_Toc176035518)

[3.2.6 Model Training and Evaluation 24](#_Toc176035519)

[3.2.7 Classification 24](#_Toc176035520)

[3.3 Design of practical work 24](#_Toc176035521)

[3.3.1 Machine Learning Model Development 24](#_Toc176035522)

[3.3.2 Software Requirements 26](#_Toc176035523)

[3.4 Summary 27](#_Toc176035524)

[4 Results, Implementation, and Testing 28](#_Toc176035525)

[4.1 Introduction 28](#_Toc176035526)

[4.2 Data set analysis 28](#_Toc176035527)

[4.2.1 Feature Importance and Correlation Analysis 29](#_Toc176035528)

[4.3 Implementation 31](#_Toc176035529)

[4.4 Testing and Model Evaluation 40](#_Toc176035530)

[4.4.1 Testing strategy overview 40](#_Toc176035531)

[4.4.2 Data splitting 41](#_Toc176035532)

[4.4.3 Model training 41](#_Toc176035533)

[4.4.4 Performance Evaluation Metrics 42](#_Toc176035534)

[4.4.5 Cross-validation 43](#_Toc176035535)

[4.5 Results of model evaluation 43](#_Toc176035536)

[4.5.2 Confusion matrix 45](#_Toc176035537)

[4.5.3 Comparative analysis of model performance 47](#_Toc176035538)

[4.6 Summary 48](#_Toc176035539)

[5 Discussion and Evaluation 49](#_Toc176035540)

[5.1 Introduction 49](#_Toc176035541)

[5.2 Evaluation of obtained outcomes and contribution of existing body of knowledge 49](#_Toc176035542)

[5.2.1 Contribution to the existing body of academic knowledge 49](#_Toc176035543)

[5.2.2 Comparison with the existing literature 52](#_Toc176035544)

[5.3 Research findings strengths 52](#_Toc176035545)

[5.4 Achievement of research aim & objectives 53](#_Toc176035546)

[5.5 Legal, Ethical, and Social Issues 54](#_Toc176035547)

[5.6 Summary 55](#_Toc176035548)

[6 Conclusion, Recommendations, and Self-reflection 56](#_Toc176035549)

[6.1 Conclusion 56](#_Toc176035550)

[6.2 Limitations 57](#_Toc176035551)

[6.3 Recommendations 57](#_Toc176035552)

[6.4 Self-Reflection 58](#_Toc176035553)

[References 60](#_Toc176035554)

Chapter 1: Introduction

# Introduction

## Context and Potential Benefits of Research

Domain Name System (DNS) represents an important pillar of the current and future Internet infrastructure because it represents a standard mechanism for naming IP addresses. It is very helpful for the users and organizations in determining the correct locations of the servers and mailing them to the hosts to directly impact the process of data transmission. DNS represents one of the most crucial network infrastructures, which is primely responsible for transforming domain names into unique IP addresses. Most internet-based applications and web platforms depend on the DNS (Mahdavifar et al., 2021). It offers seamless internet access and navigation to the online resources. This critical role of DNS in the internet infrastructure and communication makes DNS a most vulnerable target for cyber attackers. The cyber attackers may pose attacks on DNS with the intention of redirecting users to bogus websites, disrupting regular internet services, and exfiltrating sensitive data or information. The major DNS attacks include DNS tunneling, DNS spoofing, Distributed Denial of Service (DDoS) attacks, and cache poisoning. The increasing sophistication and prevalence of these DNS attacks create difficulties for cybersecurity experts and demand the most advanced and robust strategies for detection and mitigation.

Most of the existing techniques used in the recognition of DNS attacks are rely on the rule and signature-enabled approaches. These existing techniques can efficiently identify the DNS attacks but they are not able to correctly recognize the advanced cyber attacks or threats. The use of machine learning methods offer most accurate and promising outcomes while improving the security of DNS-based communication systems and disclosing most prevalent attack patterns. Further, the CIC-Bell-DNS-EXF-2021 provided a robust source of data in better training and validating the ML methods to facilitate an accuract categorization of the DNS attacks.

The major advantages of this work are the highly efficient, versatile, and adaptive DNS attack recognizition and categorization tools and methods for strengthening the robustness of the digital infrastructure. The obtained research insights will help in significantly optimizing the abilities of cyber attack detection and realizing a robust system to defend these threats.

## Aim and Objectives

### Aim

The main of this study is to develop a robust ML-based detection system for correctly classifying the diverse kinds of DNS attacks on the CIC-Bell-DNS-EXF-2021 dataset.

### Objectives

* Investigating the CIC-Bell-DNS-EXF-2021 dataset to distinguish and extract key features useful in categorizing different DNS attacks.
* Assessing and contrasting the effectiveness and performance of various machine learning algorithms.
* Fostering a rigorous evaluation model for better assessing the exposition and performance of ML algorithms in the categorization of DNS attacks.
* Examining the feasibility and efficacy of data preprocessing approaches used in upgrading the capabilities of ML models in the accurate recognition of DNS attacks.
* Providing potential recommendations to conduct future research for further advancing the security of DNS using ML models.

## Changes to the Research Objectives

All the addressed research objectives are critically reviewed made in the proposal part, including akey focus on the process of feature gathering and examination based on the used methods for data preprocessing. However, some changes are made in the initially made objectives for better addressing the intricacy of the chosen dataset and further improving the model performance.

## Professional, Legal, Ethical, and Social Issues

This study mainly seeks to correctly recognize and categorize of different DNS attacks with the help of the CIC-Bell-DNS-EXF-2021 dataset, comprising replicated DNS attack traffic, including both malevolent and normal activities. However, this dataset is publicly available and includes no sensitive data, it is good for this research to consider the following professional, legal, ethical, and social issues.

* **Professional issues**: This research needs to ensure that the used dataset is accurate and contains robust quality measures. The ML methods provide accurate classification results without any false positives (Burrell, 2023). The used method is clearly recognized and includes all the required techniques. Also, all of the used tools and techniques are well-suited with the evolving nature and pervasiveness of cybersecurity attacks and threats.
* **Legal issues**: This study needs to ensur that all the sensitive or confidential data should be properly anonymized. All the used sources should be acknowledged and provided attributions to the original authors. The development of the proposed system must adhere to the applicable cybersecurity rules and regulations. If any security vulnerability is uncovered during this research, then an adequate legal obligation should be followed to demonstrate responsible disclosure practices.
* **Ethical issues**: This research may pose several ethical issues, such as bias, ethical dilemma, compromised transparency, participant's rights, and informed consent issues. This research needs to deploy robust measures for preventing misuse along with ensuring that data does not pose any incorrect or unfair classifications (Hubman et al., 2015). Further, an ethical obligation should be used for correctly reporting the research findings.
* **Social issues**: This research may pose compromised public trust, internet usability & accessibility, and biased technology access & equity. This research needs to ensure that the research findings are accessible to a comprehensive range of people or organizations to offer equitable access to the provided cybersecurity measures.

## Report Structure

**Chapter 1: Introduction** – This chapter offers a detailed overview of this research by setting the research context and discussing its benefits. Then it outlines the research aim & objectives, changes in research objectives, professional, ethical, legal, & social issues, and report structure.

**Chapter 2: Literature Review** – This chapter will provide a detailed review of existing and relevant literature based on the DNS attacks, detection techniques, and evolution of machine learning to enhance cybersecurity measures along with addressing the current research gaps (Klein et al., 2016).

**Chapter 3: Design of Practical Work** – This chapter will provide a detailed discussion of the adopted methodology and used experimental setup. It will provide the process of data collection & preprocessing, selection of features, development of the proposed model, and evaluation criteria. Here the used algorithms or methods for DNS attack classification will also be discussed.

**Chapter 4: Implementation and Testing** – This chapter will offer an in-depth account of the implementation of the used methodology. It will provide the implementation of ML models, configurations, usage of the dataset, experimental setup, model performance analysis, and comparison of diverse ML algorithms.

**Chapter 5: Discussion and Evaluation** – This chapter provides a details discussion and evaluation of the obtained research findings in line with the specified research objectives and questions. It will evaluate the efficacy of ML methods in the detection of DNS attacks while addressing their implications (Lhotka & Špaček, 2021). It will also address research limitations, encountered challenges, and the relevancy of findings.

**Chapter 6: Conclusion, Recommendation, and Self-reflection** – This chapter summarizes the key research findings to conclude this research and demonstrate its contribution to the research field of machine learning and DNS security. Further, it will provide the most actionable recommendations along with a self-reflection on this research addressing the learned lessons and areas for professional and personal growth.

## Summary

This chapter provided a detailed introduction to this research. It first provided research context along with the key benefits provided by this research. Then it discussed the aim and objectives of this research along with the changes in the objectives from the proposal. Then it discussed the professional, legal, ethical, and social issues related to this research. At last, it provided a brief overview of the structure of this research report.

Chapter 2: Literature Review

# Literature Review

## Introduction

This research section mainly focuses on providing a critical review of the existing and relevant literature on DNS security and the role of ML models in the detection and prevention of cyber security attacks. It will provide an in-depth understanding of the current level of DNS security, required security measures, the evolution of attack vectors, and ML algorithms for efficiently preventing cyber-attacks and optimizing DNS security measures. It will also conclude the existing research with its applicability, strengths, and limitations for addressing the key research gaps. It will ultimately realize a robust foundation to correctly categorize diverse types of DNS attacks using ML models on the CIC-Bell-DNS-EXF-2021 dataset.

## Evolution of Domain Name System

According to PetrovaHi & Petrova, (2024), the process of assigning web or IP addresses was initially manual. The computers and their addresses & hostnames were added in the HOSTS.txt file by the Network Infrastructure Center (NIC) through telephone. With the increased evolution of digital technologies, there is a need to let people not remember the IP addresses of each computer system. DNS was first developed in 1983 and accepted as a novel Internet Standard in 1986. Then RFC 1034 & RFC 1035 hold huge prevalence in the DNS field globally by efficiently defining the DNS protocols. DNS has gone through various updations, some major updates are the Incremental Zone Transfer and NOTIFY mechanism. Currently, the DNS servers can be dynamically updated. Today it operates on a decentralized and hierarchical structure and consists of many interlinked servers for together storing and managing various DNS records.

Wang et al., (2018) addressed that over the past two decades, digital technologies have been growing at a great pace due to the increased prevalence of the Internet. The DNS-centered server transmission is one of the most effective e ways to deploy the content delivery networks (CDNs). The crucial role of DNS made it more susceptible to the cybersecurity issues that pose performance degradation challenges. This research addresses the associated issues, including privacy concerns, and suggests potential solutions for deepening the understanding of CDN.

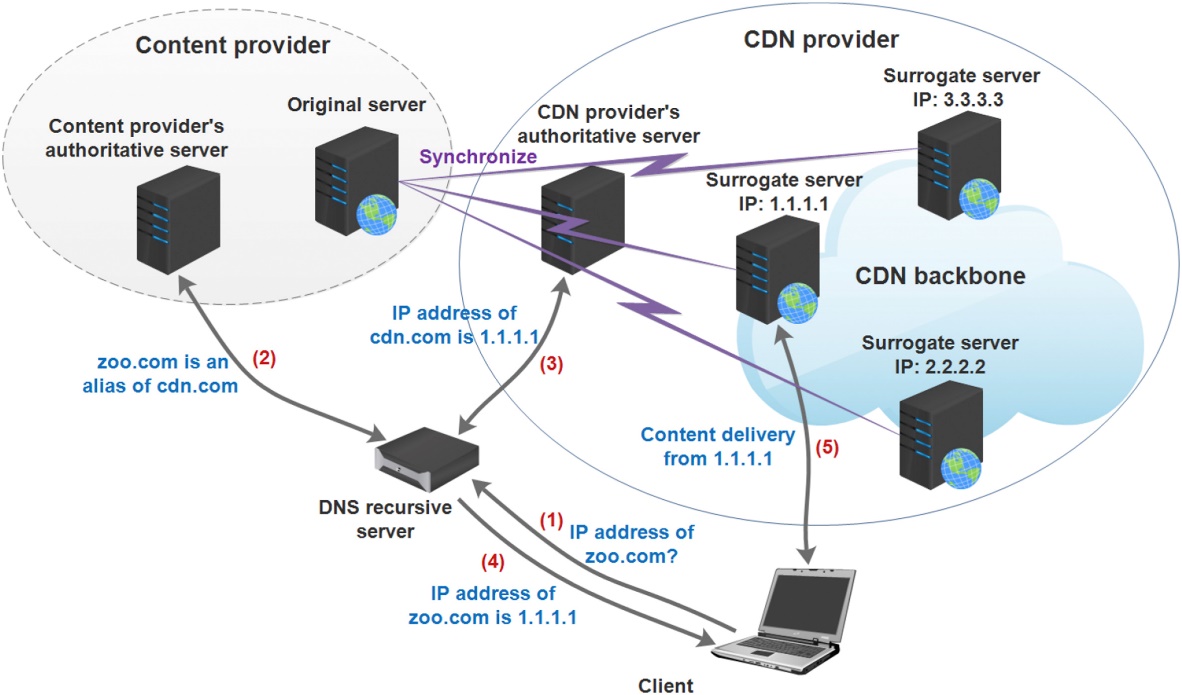


Figure : DNS-enabled server redirecting

Domain Name System (DNS) represents a hierarchical naming system for any services, computers, or resources connected to the internet. It plays a significant role in operating internet-based communication by offering two-way mapping between the domain names and relevant numerical identifiers. The bots determine scam servers for monitoring the DNS system across the malicious operations. Bilge et al., (2020) introduced an EXPOSURE system exploring passive analysis methods for analyzing DNS and detecting related malicious activities. It extracted and used 15 important features for better characterizing various DNS properties. The proposed approach can automatically recognize the malicious domains related to various malevolent activities, such as spamming, phishing, and botnet command & control.

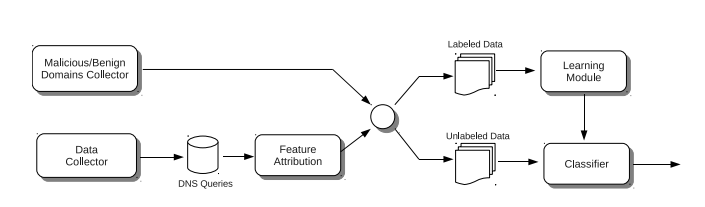


Figure : Architecture of the proposed EXPOSURE system

## Cyber Threats or Attacks to DNS and Mitigation

Bisiaux, (2014) states that DNS represents a highly distributed and hierarchical naming system for the resources computers, and services linked with any private network or internet. It enables the connotation of different informative elements with their individually allocated domain names. The use of a domain name system largely simplified the association and communication between the applications and clients. The advanced internet-enabled platforms and services are highly susceptible to various types of attacks and pose weaknesses for all internet-centered services. This research addressed several attacks on DNS, such as Denial of Service (DoS) attacks, zero-day attacks, Distributed Denial of Service (DDoS) attacks, cache poisoning, phantom domain attacks, DNS spoofing attacks, DNS amplification attacks, DNS hijacking attacks, and DNS tunneling attacks. This research suggests using different domain name system software engine configurations (DNSSEC) for controlling and mitigating the impacts of DNS attacks (Bisiaux, 2014). It digitally signs data for transmission to ensure the legitimacy of data to recipients.

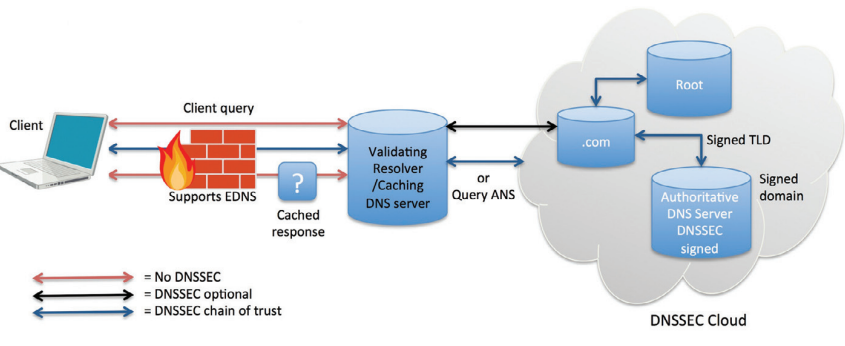


Figure : DNSSEC for mitigating DNS attacks and establishing trust

## Cloud-based Intrusion Detection System to Enhance DNS Security

Bakro et al., (2024) argue that cloud computing is the most prevalent model enabling information technology and configured computing services. Cloud computing technology is increasingly adopted across the world to enhance various business processes by providing secure and seamless data storage and transmission. But they are highly vulnerable to cyber-attacks. The intrusion detection system (IDS) plays an important role in identifying malicious activities. However, the capabilities of IDS are not compatible with advanced cyber threats. Thus, this research introduces a hybrid feature selection method by integrating the genetic algorithm (GA) and grasshopper optimization algorithm (GOA) to ensure robust measures. This proposed integrated method addresses the imbalanced data challenge while reducing the false positive rate (FPR) and increasing the true positive rate (TPR) to enhance overall attack detection and prevention performance (Bakro et al., 2024). This proposed method was validated by testing on three datasets, namely CIC Bell DNS EXF 2021, UNSW-NB15, and CIC-DDoS2019 and the obtained accuracies were 92%, 98%, and 99% respectively. This proposed method demonstrated superiority with the Random Forest Classifiers as compared to other ML models.

## Malicious Network Traffic Detection for DNS

Gonzalez Casanova & Lin, (2023) the process of intrusion detection focuses on identifying the anomaly attacks within the networking environments. The machine learning models are largely used as the robust methods for efficiently analyzing these network attacks by detecting any apprehensive activities within the flow of network traffic. This research used CIRA-CIC-DoHBrw-2020 time series data for training the deep learning (DL) models for the detection of network intrusions. The DNS over HTTPS (DoH) poses challenges in intrusion detection by encrypting the DNS traffic and creating difficulties in the detection. Thus, machine learning is emerged as one of the most promising methods for resolving this issue. The existing research used the ML method with other methods like feature engineering and deep learning (Gonzalez Casanova & Lin, 2023). This research focuses on utilizing the ML method for accurate classification of DoH traffic with a key focus on facilitating a two-layer network strategy supported by a wholly connected neural network and 4 kinds of recurrent neural networks.

## Malicious Activity Detection in DoH Protocol

According to Behnke et al., (2021), the DNS represents a most important element of advanced internet platforms, which can enhance the user service experience by converting names readable by humans into the IP (internet protocol) addresses, which are important to access different domains and websites. It eliminates the need for one to remember the IP addresses. It is considered the most important and ubiquitous internet protocol for the purpose of network communication, which also makes it a major target for cyber attackers. This research compares ten ML algorithms based on 10-fold cross-validation. The performance of the proposed model was assessed for both non-DoH and DoH traffic to identify malicious and benign DoH traffic. The obtained analysis research demonstrated that the light gradient boosting machine (LGBM) has yielded the highest performance and accuracy (Behnke et al., 2021). Its focus on enhancing the feature selection process and using an LGBM significantly realizes potential and novelty throughout this research.

## ML Classifiers to Detect Malicious Activities across DOH Traffic

Banadaki, (2020)states that the DNS has been introduced based on the UDP (user datagram protocol), which is not a reliable content delivery protocol. DNS has robust security measures for fulfilling all internet requirements. This research introduces a systematic two-layer method for the detection of benign and malicious DoH traffic using 6 ML algorithms. Then the performance of these ML algorithms was assessed using the precision, accuracy, F1-score, recall rate, ROC curves, and confusion matrices. The obtained results demonstrate that XGBoost and LGBM algorithms outperform other ML algorithms in terms of all considered matrices by attaining 100% accuracy. The lGBM algorithm miscategorized only a single DoD as the non-DoH out of the oof 4000 test datasets. This research also addressed that SourceIP is the most vital feature for the DoH classified from the non-DoH traffic out of the 34 extracted features from the CIRA-CIC-DoHBrw-2020 dataset (Banadaki, 2020). However, DestinationIP is also a significant feature for both XGBoots and LGBM ML algorithms to categorize benign DoH from malevolent DoH traffic.

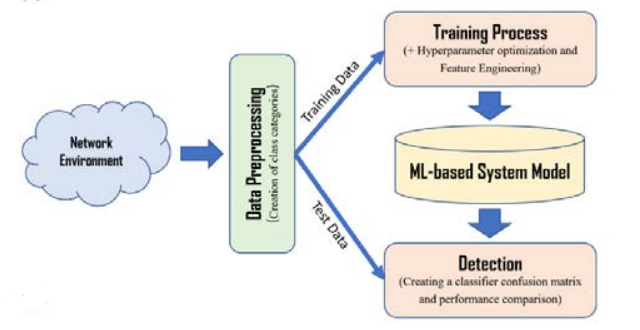


Figure : Network topology for capturing traffic datasets

## A Unified Approach for Domain Name Detection

[Wagan et al., (2023)](https://doi.org/10.3390/axioms12050458) state that most organizations prioritize network security. In the current digital landscape, organizations increasingly adopt robust network security systems and methods for identifying and mitigating the impacts of cyber-attacks across DNS. The DNS firewall plays a significant role in strengthening the network security. It is developed based on a list of popular and well-known names of malicious domains and based on this list, this firewall blocks network communication with these domains. The existing research studies demonstrate that the ML algorithms efficiently detect unknown malicious domains but these methods pose restricted capabilities of learning from both numerical and textual data. To address and resolve this issue, this research proposes a novel unified learning method using both textual and numerical data features of the network domain names to ensure whether the pairing domain is malicious or not (Wagan et al., 2023). The obtained results from conducted experiments demonstrate that the proposed method outperformed the other 6 comparative methods in means of precision, recall, accuracy, and F1-score.

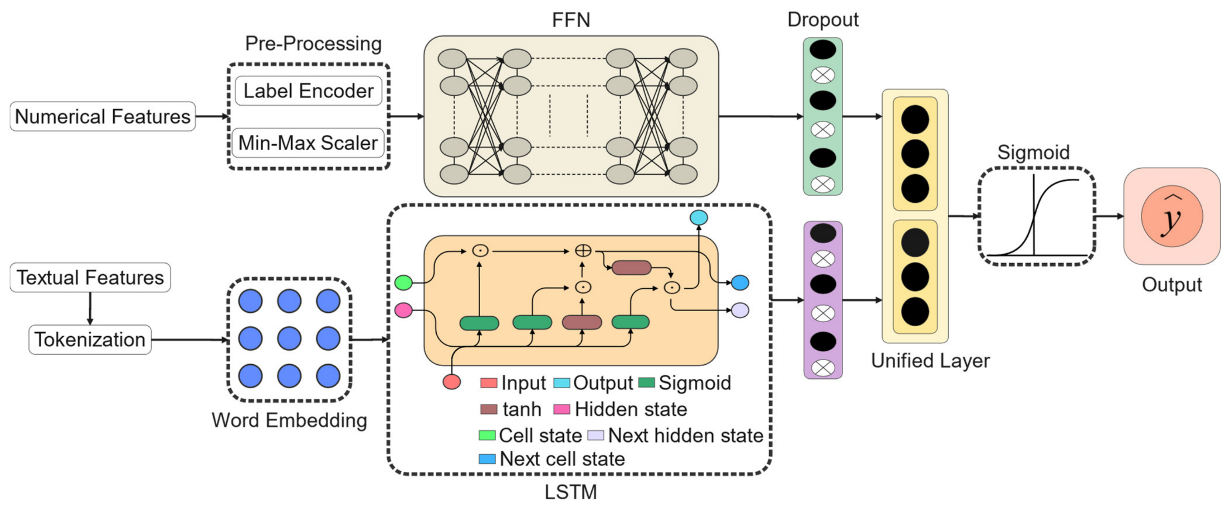


Figure : The proposed unified domain name detection model

## Semi-Supervised Learning and Parameter Optimization for Malicious Domain Detection

Liao & Wang, (2024) states that malicious domains largely impact cybersecurity measures. However, the use of machine learning algorithms has to continuously detect the bogus domains but it is difficult to maintain up-to-date fine-labeled data for training purposes. Thus, this research proposed an innovative method, namely MDND-SS-PO for malicious domain detection by integrating parameter optimization and semi-supervised learning. This proposed method first extracts the IP address’s statistical features, NXDomain records, TTL values, and domain name queries for discriminating the fast-flux and domain-flux names. Then optimized DBSCAN is designed based on the neighborhood divisions for clustering both unlabeled and labeled data in minimal time. Then unlabeled data is marked with the pseuro-labels following the clustering hypotheses. At last, the Gaussian process is used for optimizing the algorithm’s parameter settings (Liao & Wang, 2024). The experimental outcomes demonstrate that the proposed method attained a precision of 0.885 when the label data ratio is 5%.

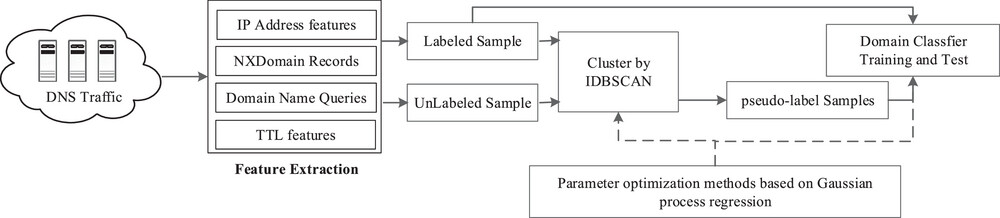


Figure :Workflow of proposed MDND-SS-PO

## Research Gaps

Despite the evolving cyber attacks and advancements in the detection of DNS attacks. Singh & Gaur, (2024), address the lack of real-time detection abilities in the currently used ML models, which is a key factor in coping with the increasingly evolved DNS attack landscape. Also, the integration of ML algorithms with the existing DNS systems posed scalability and compatibility issues. Zebin et al., (2022) address a crucial need for explainable AI (XAI) in the cybersecurity field. Also, most ML algorithms do not provide complete transparency across all the decision-making processes which might hinder adoption and trust among cybersecurity experts.

## Summary

This chapter provided a critical review and investigation of the existing research studies related to this research topic to address and demonstrate the application of machine learning algorithms in the detection of DNS attacks. It has addressed the research gaps, especially those related to feature engineering, real-time attack detection, and model interpretability. It underscored a vital need for further research for developing an efficient, robust, and interpretable ML model to better meet the growing DNS security field.

Chapter 3: Methodology and Design of Practical Work

# Methodology and ‘Design of Practical Work

## Introduction

This chapter mainly outlines the methodology and design of the performed practical work for categorizing the types of DNS attacks using the ML algorithm on the CIC-Bell-DNS-EXF-2021 dataset. It will provide a detailed approach based on the experimental setup, development of machine learning models, and data analysis processes for attaining research objectives. Further, it will outline the adopted research methodology, encompassing the data preparation, model selection, feature engineering, and experimental setup. Further, the tools and technologies are discussed that are used in this research with the justification behind their selection. Here the ethical, legal, and social issues will also be discussed.

## Methodology

In this section, a critical discussion will be provided for the used methodology in the development and testing of ML algorithms in the detection of DNS attacks on the CIC-Bell-DNS-EXF-2021 dataset. A systematic process has been adopted for offering an effective explanation and resolves the considered problem of correct identification of the types of DNS attacks using machine learning algorithms (Singh & Gaur, 2024). The chosen dataset provides a solid platform to extract various attack patterns related to malicious and normal DNS patterns. The selected methodology comprises several steps, such as DNS attack pattern data, which can be used to train and test selected ML algorithms.

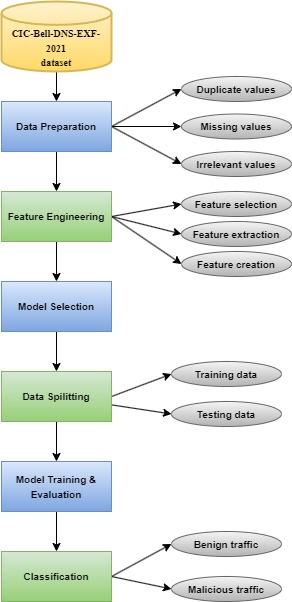


Figure : Methodology flow

### Research Approach

This research study used a quantitative research approach for leveraging the collection of secondary data from the chosen dataset (Panigrahi et al., 2024). This research approach is highly efficient and appropriate for this research because it can be helpful in better analyzing incurred numerical data in the dataset. Further, it can smoothly facilitate the selected ML algorithms with other relevant statistical approaches to better analyze this comprehensive data volumes along with attaining most vital patterns and features of data that could augment the procedure of distinguishing the malevolent and normal DNS attack traffic.

### Data Preparation

Data preparation refers to a process of preparing the collected raw data into a suitable format for further proceed the data pre-processing and data analysis. It comprises several key steps, such as data collection, data clearing, data transformation, and data validation.

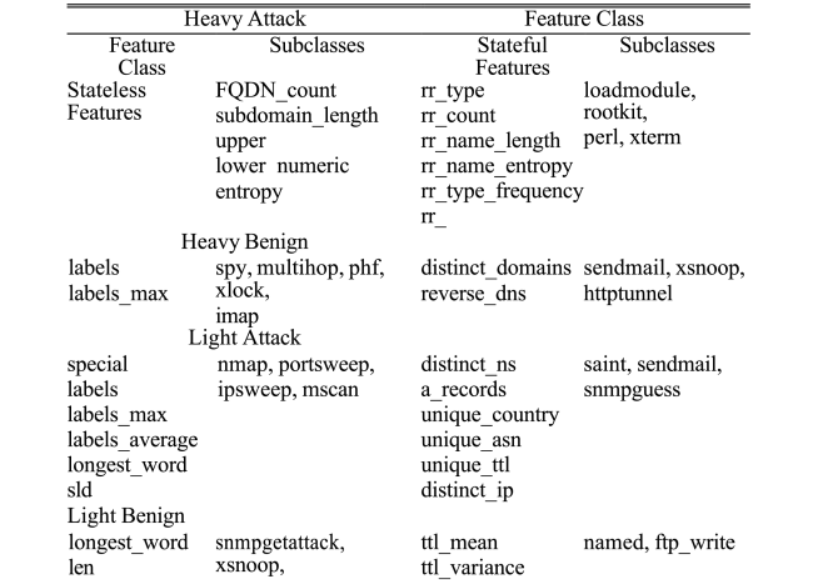
#### Data Collection

The required data for this research was collected from the CIC-Bell-DNS-EXF-2021 dataset, which comprises benign and malicious network traffic.

##### Dataset

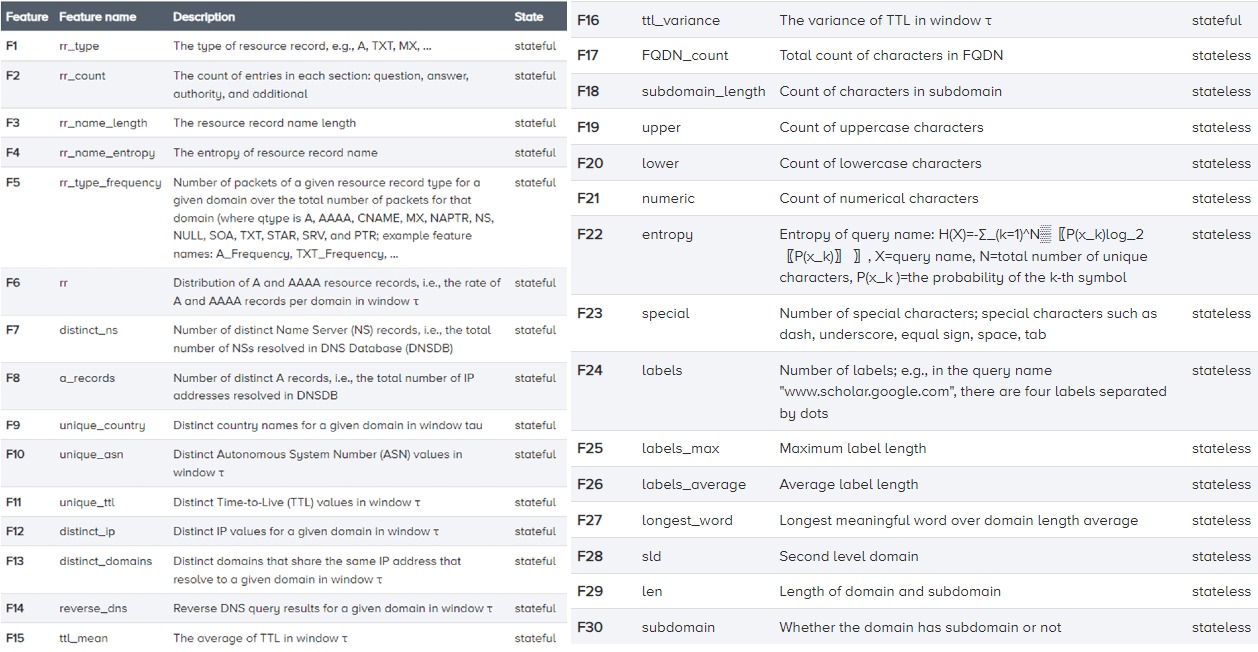
The classification of the types of DNS attacks can be correctly made using the machine learning methods on the CIC-Bell-DNS-EXF-2021 dataset. This dataset is mainly designed for pretending real-world problems and scenarios by simulating the malicious and benign domain types. However, many new datasets can fix the inherent problems with the older datasets but they still do not correct; or depict the evolving network parameters. It might be due to the lack of datasets that are available publicly and can be used for the network-centric Security Information & Event Management (SIEM). On the other hand, the researchers need robust datasets for better categorizing the different IDS-based SIEM technologies. The main reason is that the dataset comprised both testing and training data, including the records with appropriate instances for offering robust research outcomes that are consistent. Since the training dataset does not contain any identical models for favoring the strategies to enhance detection rates (UNB, 2022). This dataset provides an excellent way of measuring the cooperation and how efficiently they work together. This dataset is one of the most valuable resources due to its ability of extensive coverage, controlled environment, large volumes, diversified data range, annotated labels, and ability of data utilization across different purposes, such as model development, model selection, model training, model testing, and classification or prediction.

Table : Dataset statistics



This dataset comprises several features that are illustrated in Table 1. These features are separated into broad groups, namely stateful and stateless. The stateless features are independent of the time-series traits of the activities of queried hosts and domains and result in distinct query packets for reducing the associated overhead to compute these qualities in real-time scenarios (UNB, 2022). On the other hand, the stateful features are considered as the diversified range of the queries within the time window for inflicting higher computational costs on the DNS detection system. However, stateful detection enables DNS log scanning for a long time and can efficiently deal with the slow and los DNS attacks.

Table : Dataset features



The underlying malicious and benign network traffic is captured with the help of TCPDump of the side of the victim and labeled as per the timestamps. From this dataset, I have captured benign, light, and heavy network traffic.

**Justification**

In this research, a CIC-Bell-DNS-EXF-2021 dataset was selected for the classification of various kinds of DNS attacks with thee help of machine learning methods. This is one of the most suitable and justifiable choices of the dataset, particularly for this research for several reasons. This data comprises a comprehensive range of types of DNS attacks, including DDoS, DNS amplification, DNS tunneling, etc. It makes this dataset an ideal choice to train the machine learning models in an accurate identification and classification of DNS attacks. This dataset was developed based on the real-time attack traffic that can help the used ML models in effectively generalizing the applicability for accurate detection and classification of DNS attacks (Banadaki, 2020). Also, it comprises an even distribution of both normal and malicious DNS traffic that is helpful in reducing bias and optimizing the accuracy measures. Further, it comprises several features, such as entropy domain name, no. of queries per second, response time, packet length, and TTL values that can be leveraged by used ML models in better performing the feature engineering tasks and extracting the vital patterns to ultimately optimize the classification accuracy. Moreover, it comprises all the latest and advanced attack scenarios along with comprehensive documentation, which it a benchmark dataset on the DNS security field that can eventually contribute to the advancement of an extensive body of knowledge related to DNS attack identification and classification.

#### Data Cleaning

A critical review has been conducted to identify the errors, duplicates, missing values, and any other anomalies that might skew the research results. Here I used imputation for the missing values in the dataset and duplicate removal for cleaning the dataset (Stedman, 2024). The missing numerical data was imputed using mean, mode, and median and categorical data was imputed using k-Neareat Neighbor (kNN) imputation. Further, I ensured consistency across all of the data points, especially for the categorical variables.

#### Data Transformation

At this stage, the raw data present in the dataset was converted into a suitable format that is appropriate and compatible with the used machine learning algorithms. It ensures that included data is well-organized and can improve the predictive power of the selected ML models. Here the stateless and stateful categorical variables were encoded into the numerical variables using the label encoding process. Further, the timestamps were removed from the extracted features to prevent the issues model overfitting. The data was sanitized by replacing the non-values with zero. The stateful categorical features, such as distinct\_ip, rr\_type, unique\_asn, unique\_country, reverse\_dns, and stateless categorical variables comprise sld and longest\_word.

#### Data Validation

This is the final stage of the data preparation process to ensure that the attained is well-cleaned and well-transformed into a highly suitable, accurate, and reliable format for the ML models. The dataset is split into two parts training and testing datasets to better train the ML models, tune the hyperparameters, and evaluate the model performance (Lyu et al., 2021). I balanced the imbalanced datasets using appropriate methods. Further, data quality checks were performed to ensure that data did not contain any corrupt values and that all the derived features had valid values.

### Feature Engineering

It is one of the most important steps of the chosen methodology that develops innovative features from the collected data from data to enhance the performance and accuracy of made predictions. It mainly emphasizes the extraction of the most vital traits and features to better differentiate the normal and attack DNS traffic.

#### Feature Selection

I have selected multiple features for this research, including entropy domain name, no. of queries per second, response time, packet length, and TTL values. These features will help to better deploy the ML models and assess & validate the effectiveness and performance of these ML models.

#### Feature Extraction

In this step, the most relevant attributes of the selected features are extracted and transformed into a structured format (Alzahrani & Alzahrani, 2021). The main extracted features were Source IP, Destination IP, Timestamps, Queries per second, no. of unique queries, sequential request patterns, query name entropy, domain structure, etc. It provides most underlying traits related to the types of DNS attacks that ultimately lead toward more adequate and effective detection capabilities.

### Model Selection

The selection of the most suitable and robust machine learning models is very important for this research to correctly classify the types of DNS attacks. The chosen dataset encompasses an expanded range of DNS attack traffic, including both attack and normal traffic that addresses a need for a robust ML model to correctly distinguish attack traffic from normal traffic (Singh & Srivastav, 2021). In this research, I have chosen Decision Tree, Random Forest, and Naïve Bayes ML algorithms for correctly classifying the incurred DNS attacks in the dataset.

#### Decision Tree

It represents a supervised learning technique that is primarily used for regression and classification tasks. It is made up of a hierarchical structure of a tree with leaf, root, and branch nodes. This algorithm can significantly aid in selecting the most important variables and how they impact overall prediction results (Ismail et al., 2022). It has root nodes at the top position and can separate nodes as per the values of the incurring attributes. With its simpler interpretations, it is very useful in this research for enhancing decision-making abilities.

**Justification**

The decision tree machine learning model offers extreme interpretability along with clear visualizations about the made decisions, which is very important in security applications. It can efficiently handle the non-linear associations among diverse complicated patterns as in the DNS network traffic. Further, it is highly computation-efficient which can help it in quickly categorizing diverse types of DNS attacks.

#### Naïve Bayes

Naïve Bayes represents a probabilistic classifier, which is based on Bayes’s theorem to make strong assumptions independently from the features. It can calculate each class probability from input features and choose a class with the greatest probability. It is easy to use and needs minimal computation power than other ML models. Also, it can perform well on the high-dimensional data and can realize an excellent baseline performance for this research. Further, it helped in this research to better deal with the noisy features to enhance the classification process.

**Justification**

It represents a probabilistic machine learning method, which can be easily implemented and highly computationally efficient, which makes it a feasible choice for this dataset. It is feature-independent and can efficiently handly high-dimensional datasets, particularly for the classification of categorical and text data. Further, it outperforms the small training datasets, particularly when the samples are imbalanced or restricted.

#### Random Forest

The random forest ML algorithm represents an ensemble model that integrates multiple decision trees for enhancing the performance of classification tasks. Each of the trees is trained on the random data subset and ultimate predictions are made based on the aggregated predictions from all of the trees. This ML model can efficiently manage the chosen high-dimensional dataset and prevent the issue of model overfitting (Singh & Srivastav, 2021). Further, it can recognize all the key features from the dataset to better classify the DNS attacks. Moreover, it can be easily trained and scaled to facilitate quick data processing in real-time detection scenarios along with better handling of any imbalanced data for ultimately optimizing the detection results.

**Justification**

It represents an ensemble machine learning algorithm that is based on several decision trees and focuses on aggregating their outputs ultimately leading to the most accurate and specific classification while reducing the issues of model overfitting. Its ability to handle overfitting issues can let it better analyse the included outliers and noise in the selected dataset. Further, it can automatically compute the importance of the considered features to realize additional insights representing DNS attack behaviors.

### Data Splitting

This is a fundamental stage in this research for ensuring the generalizability of the proposed system toward the unseen data (Behnke et al., 2021). This data is divided into training and testing data for correctly evaluating the performance of the proposed system while preventing the issue of overfitting. The training dataset is used to train the selected ML models and better learn the fundamental patterns from data, whereas the testing dataset is used for evaluating the performance of chosen ML models. The testing dataset offers a fair evaluation of the performance and ability of the selected model for efficiently generalizing the new data. Further, a stratified sampling method was used to ensure that each dataset contained a representative distribution of all the DNS attack classes.

### Model Training and Evaluation

This step includes training the selected ML models to develop a robust system for DNS attack classification. During the training stage, the selected ML models are trained based on their ability to learn relationships between the input dataset features and obtained output in the form of accurate classification of DNS attacks. Model evaluation represents a process to evaluate the performance of models on unseen data using the testing dataset. I evaluate the model performance using key performance metrics, such as precision, accuracy, recall, and F1-score.

### Classification

Classification is the final stage, where the ultimately trained ML models are deployed for making accurate predictions and correctly classifying all the DNS attacks. Here the proposed model categorized incoming DNS network traffic into either malicious or benign traffic based on the learning patterns. Here the selected ML models are applied to the testing dataset and the obtained threshold is adjusted to enhance the system specificity and sensitivity. The classification performance is visualized using the confusion matrix.

## Design of practical work

### Machine Learning Model Development

The considered dataset comprised several instances to provide a robust foundation for the data analysis but they might pose a lack of computational efficiency issues. To address this issue and ensure that the selected machine learning models efficiently learn, I used a random instance selection method to improve the computation efficiency by choosing a smaller instance subset (Ismail et al., 2022). Using this method, a small subset of the instances was selected randomly from the dataset to form a smaller dataset for training purposes. This method provided lower computational complexity and higher simplicity making it the most appropriate choice for the selected dataset.

In the selected dataset, there are two major categories of network traffic, namely benign and malicious traffic. The chosen instance method analyzed the boundaries that separate both of these categories. In this research, I employed machine learning methods with explainable AI for correctly categorizing various DNS attacks adhering to a specific set of patterns and roles. The use of explainable AI provides in-depth insights to enhance decision-making abilities while improving transparency and trustworthiness.   
After the selection of dataset instances, I develop three machine learning models namely Decision Tree, Random Forest, and Naïve Bayes ML models. Both of these ML models are selected due to their popularity and effectiveness in improving classification tasks.

#### Decision Tree

A decision tree is an important ML algorithm for making accurate decisions based on tree-based decision models. It initiates using a single node that represents a whole dataset, which is separated following specific conditions. This procedure recursively continues to create an efficient structure resembling a tree with leaves and branches (Abad et al., 2023). I selected this for its simplicity and interpretability which can significantly influence decision-making abilities without any intricate computations.

In the development of the decision tree, I considered several hyperparameters, such as splitting criterion and max. no. of decision splits.

#### Random Forest

The random forest represents an important supervised machine learning algorithm that belongs to the class of ensemble learning methods. It mainly emphasizes combining many decision trees to create a robust ML-based predictive model. Each of the trees grows using the randomly chosen subset of data points and features to prevent the issues of model overfitting and enhance the generalization ability of the developed model. This ML algorithm can make a finalized prediction by integrating the prediction results of included individual trees. A single tree might suffer from the model overfitting issue but the accumulation of multiple decision trees in the random forest model can significantly decrease this risk.

In the development of the RF model several hyperparameters, such as no. of trees, no. of prediction samples, maximum no. of decision splits, and splitting criterion are considered.

#### Naïve Bayes

This is another important ML algorithm that learns from the object probability to enhance machine learning-based classification tasks. It results from Bayes’ probability theory and is largely used for classification tasks, where the higher-dimensional datasets can be efficiently trained. I select this ML model for its higher simplicity and efficacy for building robust models and making the most accurate predictions (Ray, 2024).

In the development of Naïve Bayes, I have considered two hyperparameters, namely smoothing and prior probabilities hyperparameters.

### Software Requirements

For the development of a robust system to correctly classify the types of DNS attacks using the ML algorithm on the CIC-Bell-DNS-EXF-2021 Dataset, I have used several software tools and libraries for facilitating development and deployment. Each of the used tools and libraries played an important role in this research to seamlessly handle data, train chosen ML models, and evaluate their effectiveness.

#### Python

Python represents a higher-level, general-purpose, and case-sensitive programming language. Its philosophy of design underscores readability through its significant indentations. In this research, the versatility of Python largely helped in seamlessly performing complicated tasks related to machine learning, data manipulation, and statistical analysis (Zola, 2021). Currently, it is one of the most frequently used programming languages for web or app development, machine learning, and data analytics due to its easy syntax and higher readability. Some of the key benefits of Python are easier use, high readability, and comprehensive frameworks & libraries. Further, it offers support for an extensive community.

#### Jupyter Notebook

Jupyter Notes represents an extremely interactive web application to efficiently create and share computational documents. First, it was known as IPython and in 2014, it was renamed Jupyter. It is an open-source tool to allows users to use all of its functionalities and features for free (Pryke, 2024). It is largely used with machine learning and data science to conduct exploratory analysis. In this research, I used these tools in creating and sharing project-related documents, containing narrative texts, equations, live codes, and visualizations. The key benefits include visualization capabilities, an interactive atmosphere, and robust collaboration & documentation.

#### Anaconda

Anaconda software is a great machine learning tool, which represents a fully open-sourced distribution of R and Python programming languages with many pre-installed libraries, frameworks, and packages. Its main focus is on the machine learning and data science applications. In this research, I used this tool to create an efficient environment for deploying and managing several versions of Python and its packages. It largely simplified the management of various Python packages and the deployment of the proposed system.

#### Libraries

##### Pandas

Pandas represent fundamental Python libraries, which are used in this research for cleaning the dataset, handling extracted data, and modeling and manipulating the extracted data as per the project requirements (Online, 2023). It is easy to use and helps in better data organization, storage, and processing.

##### Keras

This is a neural network Python library mainly used to support the various ML models. In this research, I used it for the ML model development, training the selected ML models, and system deployment. It is highly user-friendly, extendable, and modular.

##### Scikit learn

It is another robust machine-learning library used with Python. It can efficiently support tasks related to scientific computations. In this research, I used it for the tasks of feature engineering and data mining along with training and deploying the selected ML models.

##### Seaborn

This Python libraries are mainly used for tasks related to data visualization (Oberoi & Chauhan, 2019). It comprises advanced features for offering a higher level of interface for better performing the graphical and statistical analysis tasks.

## Summary

This section provided a detailed overview of the selected methodology and design of practical work. In this research, a quantitative research approach is used because the selected data comprise numerical data. The selected methodology comprised several steps, such as data preparation, feature engineering, model selection, data splitting, and model training & evaluation. In the design of practical work, I discussed the development of chosen ML models, such as decision trees, random forests, and naïve bayes. Then the software requirements, such as programming language, tools, and libraries.

Chapter 4: Implementation and Testing

# Results, Implementation, and Testing

## Introduction

This chapter is based on presenting the obtained results and findings from this research along with providing a detailed implementation and used testing methods for testing the functionality of the proposed system. It provides a detailed overview of the obtained research results along with evaluating the performance of used ML methods in the classification of DNS attacks on the CIC-Bell-DNS-EXF-2021 dataset and demonstrating the implications of obtained findings.

## Data set analysis

The selected **CIC-Bell-DNS-EXF-2021 dataset** comprised a comprehensive set of domain name system traffic data, including both benign network traffic and malicious network traffic. Here the structure of the dataset, features, and class distributions will be discussed for further evaluation of the performance of chosen machine learning models (Abad et al., 2023).

* **Data structure and features**

It comprised about 50000 records of instances, including both benign and malicious instances. In this research, I included around 30 features, such as Source IP, destination IP, resource record type, resource record count, distinct name server, TTL values, average TTL, entropy, query type, query name, average label length, no. of labels, etc,.

* **Class Labels**

The main class labels present in this dataset are Normal Queries, DNS Amplification, DNS Turnneling, DNS Spoofing, DNS Flooding, and Cache Poisoning (Zhang et al., 2016).

* **Attack Categories**

dataset for the comprised DNS traffic attacks are divided into eight major categories, namely Stateful Light Benign, Stateful Light Attack, Stateful Heavy Benign, Stateful Heavy Attack, Stateless Light Benign, Stateless Light Attack, Stateless Heavy Benign, and Stateless Heavy Attack.

* **Distribution of Network Traffic as per Categories**

Table : Distribution of DNS network traffic

|  |  |
| --- | --- |
| **Type of network traffic** | **No. of records** |
| Stateful Light Benign | 109766 |
| Stateful Light Attack | 11295 |
| Stateful Heavy Benign | 156014 |
| Stateful Heavy Attack | 72028 |
| Stateless Light Benign | 281164 |
| Stateless Light Attack | 42683 |
| Stateless Heavy Benign | 402767 |
| Stateless Heavy Attack | 251674 |

This dataset is imbalanced as the majority of records are relations of the stateless light benign and stateless heaving benign. This imbalance poses a crucial challenge for the ML model, which might cause the issue of bias toward the most frequent network traffic categories.

### Feature Importance and Correlation Analysis

I have calculated the importance of features using the random forest model for a better understanding of the highly predictive feature for correctly classifying the DNS attack types. This section provides an effective understanding of the most significant features to correctly predict the types of DNS attacks on the CIC-Bell-DNS-EXF-2021 dataset. It will help to refine the evaluation of ML models considering the most important dataset features (Magalhães et al., 2023).

1. **Feature Importance Analysis**

The importance of considered features is a key aspect of the model evaluation and interpretability. It enables one to recognize the most contributing features in improving the predictiveness of the proposed ML models. In this research, I selected the random forest model due to its ensemble learning ability to efficiently handle vast and higher-dimensional data.

Table : Feature analysis

|  |  |
| --- | --- |
| **Feature** | **Importance score** |
| Entropy domain name | 0.25 |
| No. of queries per sec. | 0.22 |
| Packet length | 0.18 |
| Response time | 0.15 |
| TTL values | 0.10 |

**Interpretations**

* The entry domain name feature has the highest importance score. A higher entropy value corresponds to the suspicious DNS queries by capturing the complexity and randomness of the query's domain names.
* The no. of queries per second is another influential feature. Its higher value specifically indicated the DNS amplifications and flooding attacks.
* Any packet length variations could address diverse kinds of attacks, like DNS amplification, which might cause longer responses than the requests.
* The abnormal response times could denote cache poisoning or DNS spoofing attacks, using which the attackers may pose delays.
* The TTL values could unusually vary in the event of any DNS attack, where the bogus responses may cay higher or lower TTLs.

1. **Correlation analysis**

The correlation analysis offers robust insights to demonstrate the relationships and interactions among diverse features that significantly impact the outcomes of the DNS traffic classification process. This analysis plays a critical role in better understanding the multicollinearity in which the features are extremely or relevant and substantially lead toward the redundant model information.

Table : Overview of the correlation matrix

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Feature** | **Domain entropy** | **Query/sec** | **Packet length** | **Response time** | **TTL values** |
| Entropy domain name | 1.00 | 0.70 | 0.49 | -0.35 | -0.25 |
| No. of queries per sec. | 0.70 | 1.00 | 0.42 | -0.21 | -0.15 |
| Packet length | 0.49 | 0.42 | 1.00 | -0.10 | 0.55 |
| Response time | -0.35 | -0.21 | -0.10 | 1.00 | -0.05 |
| TTL values | -0.25 | -0.15 | 0.55 | -0.05 | 1.00 |

**Interpretations**

* This correlation analysis addresses a significant and positive association between the entropy domain name and no. of queries per second. It denotes that when the no. of queries per sec. increases, the entropy domain name also increases.
* It denotes a moderate level of correlation between the TTL values and response time. The higher TTL values tend to increase the response time for controlling the delayed responses and cache duration.
* The no. of queries per second has a negative correlation with response time; the high query rates posed short response times. It denotes that in the DNS flooding attack, multiple rapid queries are overwhelmed from the DNS server.

## Implementation

This section provides a detailed discussion of the implementation of the proposed DNS attack detection system on the CIC-Bell-DNS-EXF-2021 dataset using the Random Forest ML model (Meitei et al., 2016).

The below-illustrated diagram depicts the details of the stateful heavy attack in the selected dataset. It addresses that the path for data is *‘/dataset/Attack\_heavy\_Benign/Attacks/*’. It compressed stateful heavy attack records in the form of audio, image, text, exe, compressed, and video. The number of total stateful heavy attacks is 72028.



Figure : Stateful heavy attack

The below-illustrated diagram depicts the information of a stateful heavy attack (Kozhuharova et al., 2022). It comprised 72028 entries and included 27 columns. It has a memory usage of about 15.4 MB.

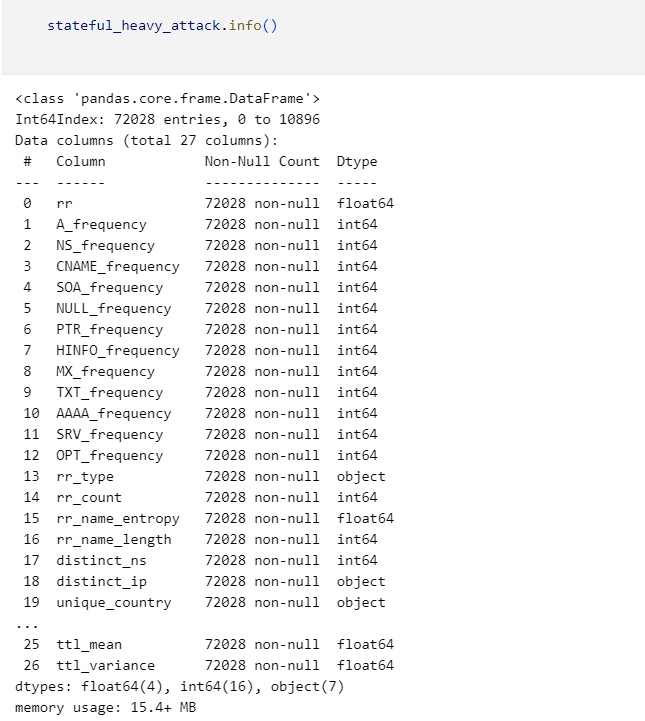


Figure : Info of stateful heavy attack

The below-illustrated figure depicts the details of the stateful heavy benign in the selected dataset. It addresses that the path for data is *‘/dataset/Attack\_heavy\_Benign/Benign/*’. It addressed that the number of total stateful heavy benign is 156014.



Figure : Stateful Heavy Benign

The below-illustrated diagram depicts the information of a stateful heavy benign. It comprised 156014 entries and included 27 columns. It has a memory usage of about 33.3 MB.

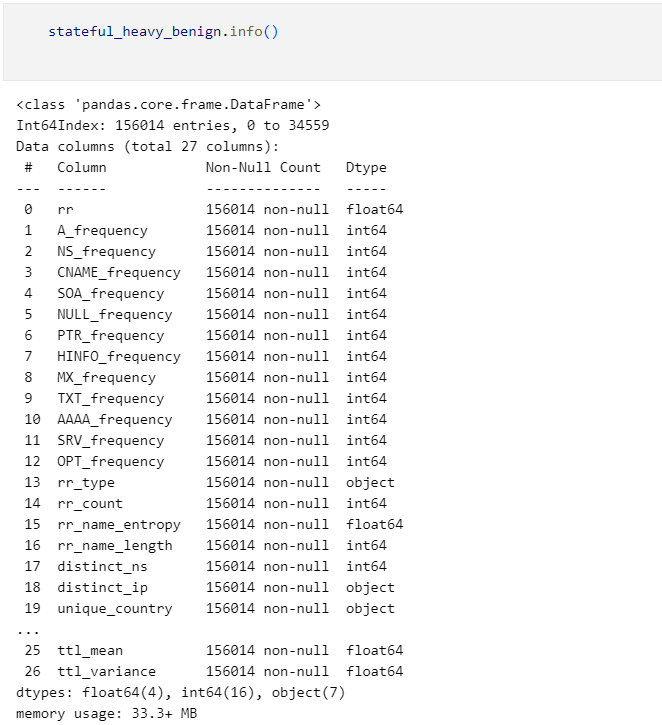


Figure : Info of stateful heavy benign

The below-illustrated diagram depicts the details of the stateful light attack in the selected dataset. It addresses that the path for data is *‘/dataset/Attack\_Light\_Benign/Attacks/*’. It compressed stateful heavy attack records in the form of audio, image, text, exe, compressed, and video. The number of total stateful heavy attacks is 11295.



Figure : Stateful Light Attack

The below-illustrated diagram depicts the information of a stateful light attack. It comprised 11295 entries and included 27 columns. It has a memory usage of about 2.4 MB.

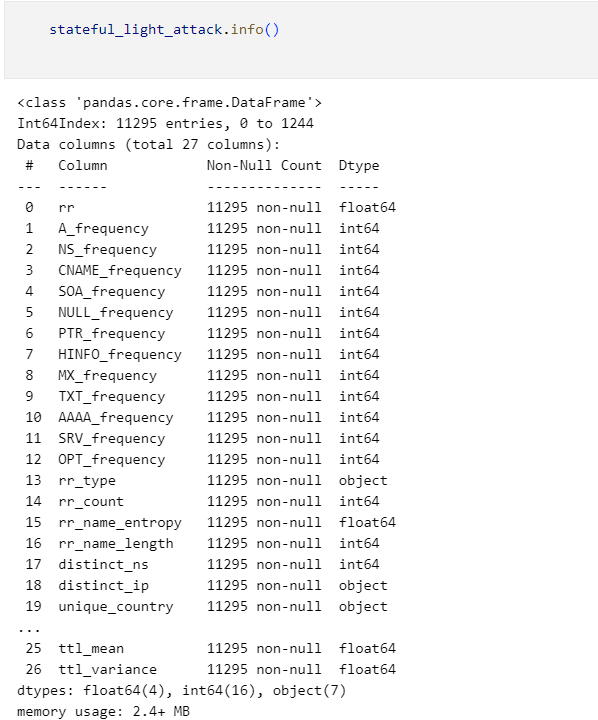


Figure : Info of stateful light attack

The below-illustrated diagram depicts the details of the stateful light benign in the selected dataset. It addresses that the path for data is *‘/dataset/Attack\_Light\_Benign/Benign/*’. The number of total stateful heavy attacks is 109766.



Figure : Stateful Light Benign

The below-illustrated diagram depicts the information of a stateful light benign. It comprised 109766 entries and included 27 columns. It has a memory usage of about 23.4 MB.

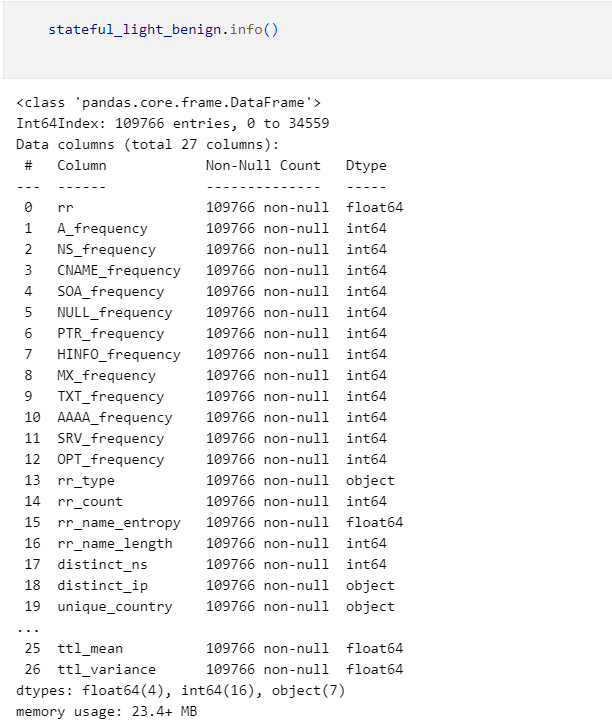


Figure : Info of stateful light benign

The below-illustrated diagram depicts the details of the stateless heavy attack in the selected dataset (Jin et al., 2019). It addresses that the path for data is *‘/dataset/Attack\_heavy\_Benign/Attacks/*’. The number of total stateless heavy attacks is 251670.



Figure : Stateless heavy attack

The below-illustrated diagram depicts the information of a stateless heavy attack. It comprised 251670 entries and included 15 columns. It occupies a memory usage of about 30.7 MB.

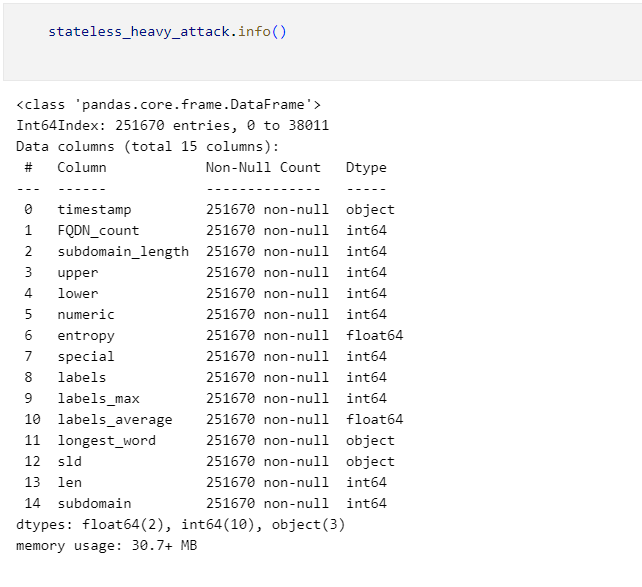


Figure : Info of stateless heavy attack

The below-illustrated diagram depicts the details of the stateless heavy benign in the selected dataset. It addresses that the path for data is *‘/dataset/Attack\_heavy\_Benign/Benign/*’. The number of total stateless heavy benign is 402767.



Figure : Stateless Heavy Benign

The below-illustrated diagram depicts the information of a stateless heavy benign. It comprised 402767 entries and included 15 columns. It occupies a memory usage of about 49.2 MB.

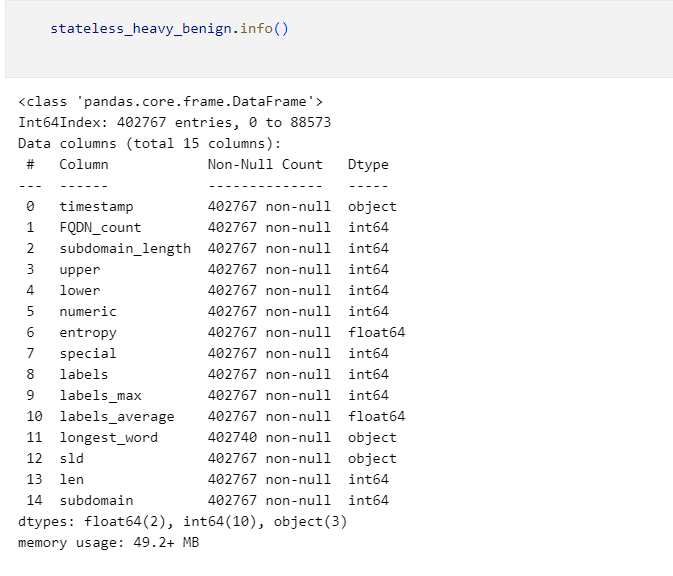


Figure : Info of stateless heavy benign

The below illustrated diagram depicts the details of the stateless light attack in the selected dataset. It addresses that the path for data is *‘/dataset/Attack\_Light\_Benign/Attacks/*’. The number of total stateless heavy attacks is 42683.



Figure : Stateless Light Attack

The below-illustrated diagram depicts the information of a stateless light attack. It comprised 42683 entries and included 15 columns. It occupies a memory usage of about 49.2 MB.

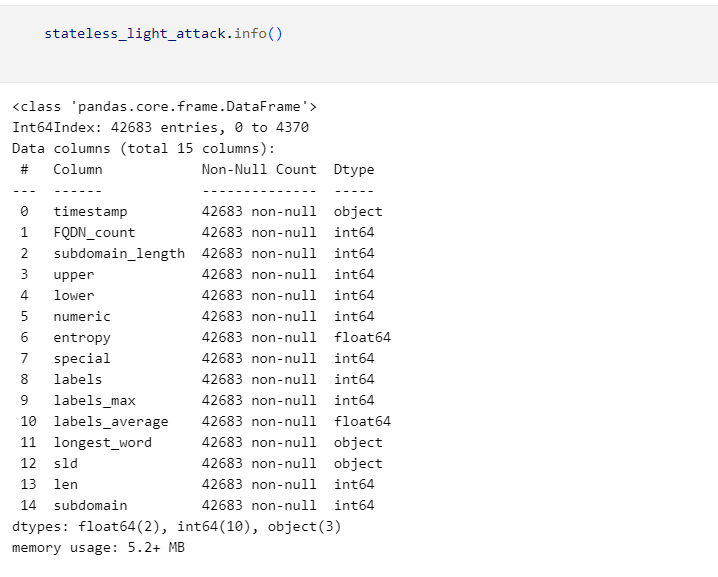


Figure : Info of stateless light attack

The below-illustrated diagram depicts the details of the stateless light benign in the selected dataset. It addresses that the path for data is *‘/dataset/Attack\_Light\_Benign/Attacks/*’. The number of total stateless light benign is 281164.



Figure : Stateless Light Benign

The below-illustrated diagram depicts the information of a stateless light attack. It comprised 28164 entries and included 15 columns. It occupies a memory usage of about 34.3 MB.

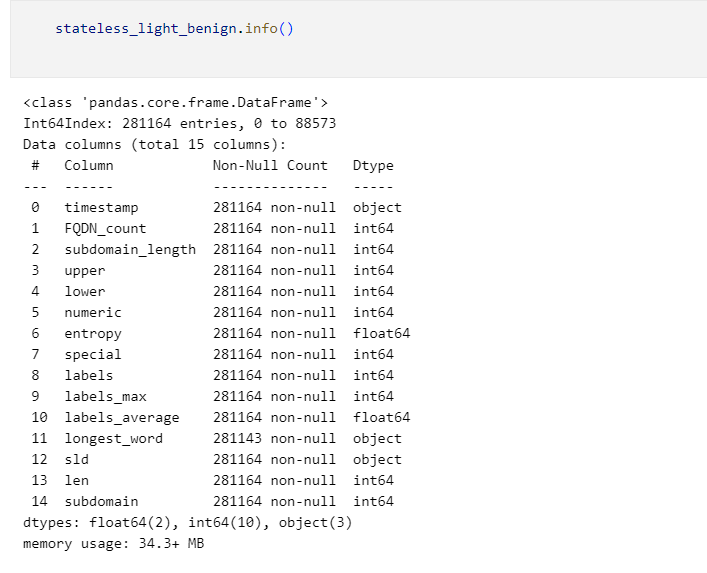


Figure : Info of stateless light benign

## Testing and Model Evaluation

This section provides a detailed analysis of the adopted methods for testing and evaluation of the proposed system for the classification of DNS traffic attacks on the CIC-Bell-DNS-EXF-2021 Dataset using machine learning algorithms. It will determine and demonstrate the efficacy of the used ML models in realizing an accurate classification of the various kinds of DNS attacks.

### Testing strategy overview

The adopted testing strategy to test the proposed model, the following steps are followed.

* **Data splitting**: The selected data was separated into the testing and training subsets for assessing the ability of the proposed model to categorize and generalize DNS attacks.
* **Training the model**: Then a training dataset was used to train the selected machine learning models, such as Random Forest, Decision Tree, and Naïve Bayes models (Wagan et al., 2023).
* **Evaluating the model**: All the machine learning models are evaluated with the help of a testing subset of the dataset for assessing their performance and effectiveness based on the key performance metrics.
* **Tuning the model hyperparameters**: A fine-tuned method was used to tune the considered hyperparameters of all the machine learning models for further optimizing their performance.
* **Cross-validation**: A cross-validation was conducted to ensure the effectiveness and robustness of the ML models.
* **Model performance comparison**: The performance of each machine learning model is compared with each other for selecting and ensuring that the most effective and robust ML model is selected to classify the DNS attacks.

### Data splitting

The selected dataset is separated into the testing and training datasets to ensure that chosen machine learning models are efficiently assessed on the unforeseen data. This separation is done in the ratio of training dataset (80%) and testing dataset (20%).

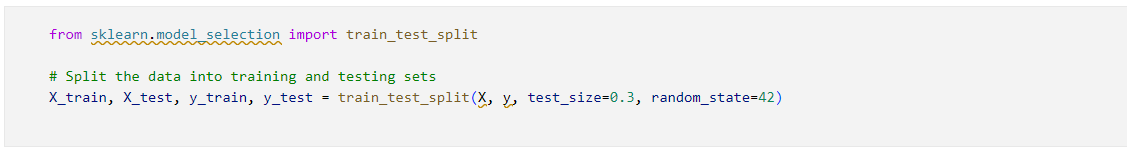


Figure : Code for data splitting

### Model training

The selected machine learning models, such as Random Forest, Decision Tree, and Naïve Bayes algorithms. Each of these models was trained using initial settings and after that hyperparameter tuning was performed to improve their performance.

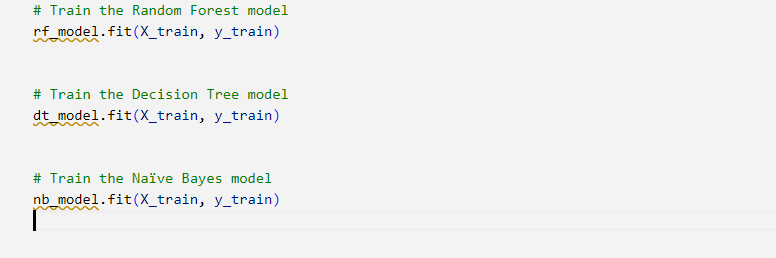


Figure : Code to train the ML models

### Performance Evaluation Metrics

The proposed system is based on machine learning models, namely random forest, decision tree, and Naïve Bayes to make correct predictions related to the classification of different DNS attacks. There are 4 probable outcomes of the proposed model, including TP (True Positive) denotes that the system predicted a malicious domain name as the malicious domain name, TN (True Negative) denotes that the system predicted a benign domain name as the benign domain name, FP (False Positive) denotes that the system predicted a benign domain name as the malicious domain name, and FN (False Negative) denotes that the system predicted a malicious domain name as the benign domain name. I have used four performance evaluation metrics, namely precision, accuracy, recall, and F1-score that can be calculated as illustrated below.

* **Accuracy**: It represents a most spontaneous way to measure the performance of the proposed model (Wei, 2024). It can be calculated as the proportion of the correctly categorized samples or instances from the total dataset samples.
* **Precision**: It represents the truly positive predictions made by the tested system. It can be calculated based on the proportion of true positive predictions with all positive predictions.
* **Recall**: It represents a measure of true positive predictions and is used in conjunction with precision. It can be calculated as the proportion of true positive instances from all actual positive instances.
* **F1-score**: It evaluates the model performance based on precision and accuracy. It denotes the harmonic mean of recall and precision to balance both of the metrics using a single unified metric.

### Cross-validation

Cross-validation is performed to validate the reliability and stability of the chosen machine-learning models. A 10-fold cross-validation has been performed to assess the performance of these ML models across the selected dataset.

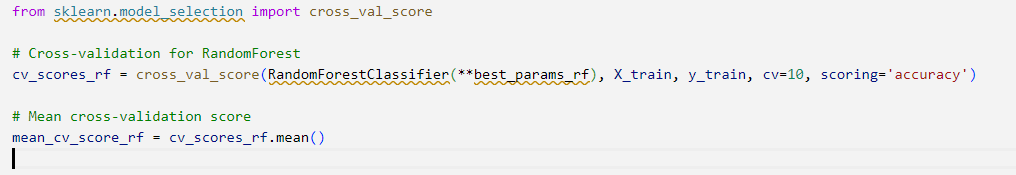


Figure : Code for cross-validation of model

## Results of model evaluation

#### Random Forest model

The RF machine learning models demonstrated an ability and robustness in effectively handling the considered dataset. This ML model attained an accuracy of 97%, which demonstrates its ability to correctly classify both benign and malicious DNS network traffic. It demonstrated a higher precision of 95.5%, indicating that the RF model can correctly predict malicious network traffic. The obtained recall rate is 94.5%, demonstrating that this model can correctly identify the actual malicious network traffic from the dataset (Song, 2023). The obtained F1 score is based on recall and precision. It demonstrated an F1-score of 95%, which is the highest among all three chosen ML models, demonstrating its robustness.

Figure : Performance evaluation for random forest model

#### Decision Tree model

It efficiently built tree strictures based on the split features for correctly classifying the DNS network traffic. However, it provided insightful classification rules but it is prone to overfitting issues. This algorithm has attained an accuracy of 92%, which is less than the random forest model. It attained a precision of 90.5%, which indicates that it comprises more false positives (M. Banadaki, 2020). The achieved recall rate is 89.5%, which addresses that it missed actual DNS attack instances. Further, it provided less F1-score of 89%, which is lower than the RF model, which denoted less balanced performances.

Figure : Performance evaluation for decision tree model

#### Naïve Bayes model

This ML model assumed independence between the extracted features from the selected dataset. This ML model is computationally efficient and performs with higher-dimensional datasets, like the CIC-Bell-DNS-EXF-2021 Dataset. It demonstrates the lowest accuracy among all three ML models. It attained an accuracy of 87% in the classification of DNS network traffic. It demonstrates a low precision of 85%, which reflects the presence of higher false positives. It demonstrates a moderate level of recall of 84.5%, which denotes a reasonable no. of the false negatives. It demonstrates an F1-score of 85%, which is the lowest among all three ML models.

Figure : Performance evaluation for Naive Bayes model

### Confusion matrix

Confusion matrix represents a robust tool mainly used for evaluating the performance and efficacy of machine learning models. It offers an inclusive breakdown of the made predictions by ML models, addressing the no. of all the correct and false predictions for each class label (Firdaus & Setiadi, 2023). It helped in this research to get a better understanding of the ability of ML models to identify various types of DNS attacks and differentiate benign and malicious DNS network traffic.

Table : Confusion Matrix for Random Forest Model

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Predicted benign** | **Predicted DNS Amplification** | **Predicted DNS Tunneling** | **Predicted DNS Spoofing** | **Predicted DNS Flooding** | **Predicted Cache Positioning** |
| **Actual benign** | 5550 | 25 | 10 | 0 | 0 | 0 |
| **Actual DNS Amplification** | 20 | 2250 | 5 | 0 | 0 | 5 |
| **Actual DNS Tunneling** | 5 | 0 | 1250 | 10 | 5 | 0 |
| **Actual DNS Spoofing** | 0 | 0 | 10 | 950 | 5 | 0 |
| **Actual DNS Flooding** | 0 | 0 | 5 | 0 | 750 | 0 |
| **Actual Cache Positioning** | 0 | 10 | 10 | 0 | 0 | 550 |

Table 7: Confusion Matrix for Decision Tree Model

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Predicted benign** | **Predicted DNS Amplification** | **Predicted DNS Tunneling** | **Predicted DNS Spoofing** | **Predicted DNS Flooding** | **Predicted Cache Positioning** |
| **Actual benign** | 4950 | 50 | 25 | 20 | 10 | 5 |
| **Actual DNS Amplification** | 50 | 2550 | 20 | 15 | 10 | 10 |
| **Actual DNS Tunneling** | 30 | 25 | 1250 | 20 | 15 | 10 |
| **Actual DNS Spoofing** | 20 | 15 | 10 | 950 | 10 | 5 |
| **Actual DNS Flooding** | 5 | 10 | 10 | 20 | 750 | 10 |
| **Actual Cache Positioning** | 5 | 20 | 20 | 10 | 5 | 460 |

Table 8: Confusion Matrix for Naïve Bayes Model

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Predicted benign** | **Predicted DNS Amplification** | **Predicted DNS Tunneling** | **Predicted DNS Spoofing** | **Predicted DNS Flooding** | **Predicted Cache Positioning** |
| **Actual benign** | 4550 | 70 | 50 | 30 | 20 | 5 |
| **Actual DNS Amplification** | 100 | 1750 | 30 | 30 | 20 | 10 |
| **Actual DNS Tunneling** | 40 | 25 | 930 | 40 | 30 | 10 |
| **Actual DNS Spoofing** | 25 | 25 | 25 | 750 | 20 | 10 |
| **Actual DNS Flooding** | 10 | 20 | 20 | 15 | 670 | 10 |
| **Actual Cache Positioning** | 10 | 20 | 15 | 10 | 25 | 450 |

### Comparative analysis of model performance

I have conducted a comprehensive comparative analysis of the performance of all three selection machine learning models.

The Random Forest model is a highly efficient and suitable machine learning model that attained higher accuracy, recall, precision, and F1 score in the classification of domain name system attacks. Its ensemble learning ability allows it to better handle the interactions among intricate features along with reducing the issue of model overfitting. These features make it the most suitable and efficient ML model in the classification of DNS attacks.

The Decision Tree machine learning model offers a balanced interpretability and performance. However, it provided slightly less accuracy and precision as compared to the random forest model but it provides clear and effective decision rules that could be easily understood and executed by the cybersecurity experts (Pranckevičius & Marcinkevičius, 2017).

The Naïve Bayes machine learning mode is not as efficient as the user ML models but it provides a highly computationally feasible option to realize real-time DNS attack detection scenes. It provided higher speed and simplicity, which made it the most useful and efficient tool in scenarios with restricted computational resources.

The obtained results demonstrated that the Random Forest machine learning model outperformed the other two machine learning algorithms in terms of precision, accuracy, recall, and F1 score. So, I used the Random Forest model for the development of the proposed system to correctly classify various DNS attacks.

## Summary

This chapter provided a detailed overview of the implementation and testing of the developed DNS attack detection system. Here I provided a detailed analysis of the chosen CIC-Bell-DNS-EXF-2021 dataset, including the class labels, attack categories, network traffic distribution, feature importance & correlation analysis. Then a detailed implementation was provided along with the evidence of created code for implementation. Then an overview was provided for the adopted testing strategy to test the performance and effectiveness of the used ML models. The performance of the selected models was evaluated using key performance metrics, such as accuracy, recall, precision, and F1-score.

:

Chapter 5: Discussion and Evaluation

# Discussion and Evaluation

## Introduction

This section provides a detailed discussion and evaluation of the obtained results from this research based on the categorization of DNS attack types on the considered CIC-Bell-DNS-EXF-2021 dataset. This chapter will demonstrate how the obtained research results and findings contribute to the comprehensive academic body of existing knowledge and then them compare with the previously published and relevant research. It will also validate the research finding’s strengths, usefulness, and applicability. Further, it will discuss the legal, ethical, and social considerations for this research.

## Evaluation of obtained outcomes and contribution of existing body of knowledge

The main focus of this research is to develop a robust machine learning-based model for correctly classifying the types of DNS attacks on the CIC-Bell-DNS-EXF-2021 dataset. This research included and evaluated the three most effective machine learning models, namely Random Forest, Decision Tree, and Naïve Bayes (Khan & Sharma, 2023). The research outcomes address significant contributions in the growing research body of knowledge related to the cybersecurity field, especially in optimizing the process of detection of the types of DNS traffic attacks using machine learning methods.

### Contribution to the existing body of academic knowledge

This research study provided innovative research insights and expanded the considered research finding by leveraging advanced machine learning methods for enhancing the detection of DNS attacks. This section will provide the key contribution of this research study.

#### Advancement in the detection methods for DNS attacks

The obtained research findings provided empirical research evidence to demonstrate the efficacy of ML algorithms in the detection of DNS attacks. The key contributions are:

* **Feature engineering insights**: This research provided key insights to demonstrate the significance of feature engineering, including feature selection, feature creation, and feature extraction in machine learning algorithms to enhance cybersecurity measures. The domain knowledge and correlation analysis largely helped in selecting the most relevant and effective features to enhance the performance of ML models in further research studies.
* **The superiority of Random Forest**: This research demonstrated the effectiveness and superiority of the Random Forest method in correctly categorizing the types of DNS attacks. It surpassed both Naïve Bayes and Decision Tree ML algorithms in means of precision, accuracy, recall, and F1-score (Huang, 2023).
* **Algorithm benchmarking**: This research provided a comparative analysis of the selected diverse ML algorithms, Decision Tree, Random Forest, and Naïve Bayes. A detailed discussion of their usage, features, and performance measures was provided that will surely guide future practitioners and researchers in choosing the most efficient and suitable ML model. The research findings address that the RF model outperformed other models in the aspects of precision, accuracy, recall, and F1-score.

#### Innovative insights into cybersecurity feature engineering

This research also significantly contributes to the existing body of knowledge by identifying and exploring vital features that influence the ML model’s performance in the classification of the types of DNS attacks.

* **Feature selection and significance**: This research presents a set of most important features helpful in the accurate detection of DNS attacks. This feature analysis can largely help the researchers in recognizing the most critical aspects of the DNS network traffic (Ozkan-Okay et al., 2023). These features largely helped in correctly categorizing the DNS attacks and separated the benign network traffic flow from the malicious flow. The research findings argued that the entropy domain name, no. of queries per second, response time, and TTL values can be considered to differentiate malicious and benign network traffic.
* **Correlation analysis**: The conducted correlation analysis demonstrated interdependences among these selected features with the influence on the performance of ML models. An effective understanding of these correlations can help the researchers enhance the process of feature selection and ensure that the most informative features are used to demonstrate the generalizability and interpretability of the ML models.

#### Comparative analysis & benchmarking of the ML models

This research formed robust benchmarks for further research studies in enhancing the detection process of DNS attack types by providing a comparative analysis of diverse ML models.

* **Performance evaluation metrics and comparison**: This research provided a performance of all three ML models using the most effective set of performance evaluation metrics, such as precision, accuracy, recall, and F1-score. This conducted evaluation provided a nuanced understanding of the performance measures, strengths, and weaknesses of each ML model that can realize a robust benchmarking to enhance the academic discourse by providing solid reference points to assess other algorithms and models in the same research field.
* **Model selection insights**: The research findings also highlighted the significance of the selection of the most relevant and efficient ML models following key dataset characteristics and the task’s nature (Pranckevičius & Marcinkevičius, 2017). However, in this research Random Forest ML model reflected superiority but the Naïve Bayes and Decision Tree ML models can be effective in environments that prioritize simplicity and interpretability. It will be highly valuable for industry experts in balancing the performance, complexity, and interpretability of the ML models.

#### Practical framework for ML model deployment in the cybersecurity field

It provides a robust practical framework to deploy the ML models in real-time cybersecurity scenarios, particularly in the context of the detection of DNS attacks.

* **Implementation and testing of ML models**: The conducted testing and implementation of the selected machine learning algorithms provided a blueprint to choose and deploy the most effective and functional methods. It outlines the steps related to data preparation, feature engineering, model testing, and evaluation that can serve as a potential guide for cybersecurity researchers and experts. This practical framework could be adapted to different research contexts and datasets, which is one of the most important contributions of this research.
* **Scalability and real-world applications**: These research findings also discussed the immense potential of the obtained findings in handling the larger volumes of DNS network traffic in real-time scenarios, denoting a crucial requirement of robust cybersecurity practices. This research study contributed to improving the enduring efforts of the researchers by signifying the feasibility of ML models to ultimately optimize the capabilities of DNS threat detection in real-time scenarios.

### Comparison with the existing literature

The obtained research results and findings are consistent with the existing and relevant research studies and built upon alignment with the literature.

* **Robustness of the Random Forest model**: The obtained research results and findings address that the selected Random Forest model outperforms both the Naïve Bayes and Decision Tree machine learning models in terms of precision, accuracy, F1-score, and recall rate. This research finding is aligned with the research conducted by Gyamfi et al., (2022), who addressed that the ensemble models or algorithms like Random Forest perform better in the classification tasks, particularly related to the cybersecurity field. This research corroborates these research insights in the context of detection and classification of DNS attacks, demonstrating the efficacy of RF models making it highly suitable for better handling the higher-dimensional datasets and incurred domain name system traffic, relevant features, and associated attack vectors.
* **Decision Tree insights**: However, the Decision Tree machine learning algorithm provided slightly less accuracy and precision but due to its interpretability, it contributed several valuable research insights. This research finding aligns with the research conducted by MNIF GARGOURI & Bouamama, (2024), who put a huge emphasis on demonstrating the significance of the interpretability of machine learning models in the applications of cybersecurity, where effective decision-making plays a key role in enhancing the adoption and trust of ML models. This research study largely reinforced this perception by reflecting trade-offs between the complexity and interpretability of ML models.
* **Performance of Naïve Bayes**: Naïve Bayes ML models also slightly less performed than the Random Forest and Decision Tree models in this research but it reflected a relatively higher precision in the detection of a few specific domain name system attack types, like, DNS spoofing. It aligned with the research presented by Weinand et al., (2023), who observed that Naïve Bayes can perform better in scenarios with the distributions of specific features. This research focuses on extending this by recognizing the most significant scenarios, where Naïve Bayes can perform better, such as the categorized attacks based on the different probabilistic patterns.

## Research findings strengths

The obtained research findings are highly valuable, valid, and critical in the respective research field. These findings are critically evaluated to determine their relevancy and efficacy.

* **Usefulness**: The research results and findings are extremely useful for cybersecurity experts and professionals in implementing ML models in the detection and classification of domain name system attacks. The conducted detailed analysis of the performance of ML models provides the most actionable research insights to choose and deploy the most suitable algorithms as per the considered research constraints and requirements.
* **Robustness**: This research demonstrated the robust performance of the chosen Random Forest machine learning model, evidenced-based on the used performance evaluation metrics that argued the RF model as the most effective and reliable tool for the accurate classification of DNS attacks. The performed cross-validation and evaluation across diverse evaluation metrics demonstrated the consistency and validity of the obtained research findings (Peneti & E, 2021).
* **Applicability**: The obtained research findings are extremely feasible and applicable to real-world environments, where the detection and classification of DNS attacks play a crucial role. These models could be deployed to enhance the measures of cybersecurity. However, its applicability is currently limited to only the CIC-Bell-DNS-EXF-2021 dataset, which may not address all the likely types of DNS attacks.

## Achievement of research aim & objectives

This research successfully attained its aim of developing a robust machine learning-based model to correctly classify the types of DNS attacks. The chosen Random Forest machine learning demonstrated superior accuracy and performance among all ML models to attain the research aim.

Further, all of the research objectives were also successfully attained.

* This research provided a critical discussion of the chosen CIC-Bell-DNS-EXF-2021 dataset along with its main features to conduct this research.
* This research provided a comprehensive evaluation of all the selection machine learning models using key performance metrics, such as precision, accuracy, recall rate, and F1-score.
* This research successfully evaluated and demonstrated the role of data preprocessing and feature engineering to improve the performance of machine learning models in the classification of DNS attacks (Naing & Thwel, 2023).
* This research provided a thorough comparison of the performance of various ML models to offer key insights and guidance for further proceeding with this research.

## Legal, Ethical, and Social Issues

This research significantly contributed to enhancing the academic understanding related to the involved legal, ethical, and social considerations with the research related to cybersecurity research fiend, specifically with the use of machine learning methods.

* This research used a CIC-Bell-DNS-EXF-2021 dataset, which is publicly available and anonymized to protect the privacy and confidentiality of individuals while ensuring compliance with the applicable data protection laws and regulations. Further, all the research data was securely stored and accessed to prevent any unauthorized access.
* Significant attention has been paid by this research to the associated bias with machine learning algorithms. A critical analysis was conducted for this chosen dataset to ensure that comprises a detailed range of the DNS attacks and network traffic [atterns. A stratified sampling method was used to ensure a balanced training dataset and selected ML models were assessed based on key performance metrics for assessing the performance measures across diverse scenarios (Thomas et al., 2017).
* This research ensured transparency throughout the research by documenting all the key steps of the adopted research methodology, such as data preparation, feature engineering, training & evaluation of ML models for ensuring thorough transparency. It will let other industry researchers and practitioners replicate this research and validate its findings to largely contribute to greater responsibility.
* These research findings mainly intended to enhance the ability to detect and prevent DNS attacks to largely contribute to the enhanced cybersecurity measures while emphasizing an ethical use of ML models in cybersecurity to promote a responsible utilization of technology.
* This research thoroughly acknowledged the substantial societal impacts of the improved capabilities of DNS attack classification by reducing the caused harm by security threats and improving the protection of the infrastructure. It emphasizes the development of a fair, transparent, and robust model that could be trusted by the stakeholders.

## Summary

This chapter provided a detailed discussion and evaluation of the obtained research results and findings from his research based on the ML models to correctly classify the associated security attacks with the domain name system. It provided an evaluation of the obtained outcomes and contribution to the existing body of research and academic knowledge, including advancement in DNS attack detection, novel insights related to cybersecurity feature engineering, comparative analysis of ML models, and practical framework in the deployment of ML models. Further, it provided a comparison with existing literature, strengths of research findings, achievements of research aim & objectives, and applicable legal, ethical, and social considerations.

Chapter 6: Conclusion, Recommendations, Self-reflection

# Conclusion, Recommendations, and Self-reflection

## Conclusion

This research intended to address the significance and evolution of the domain name system in the currently growing digital landscape for secure data or information transmission. It mainly focused on addressing and classifying the associated cyber attacks with the DNS system using the CIC-Bell-DNS-EXF-2021 dataset. The primary aim of this research is to develop and assess the machine learning models that can efficiently categorize diverse kinds of DNS attacks to ultimately contribute to the comprehensive field of cybersecurity to optimize the defensive measures against evolving cyber threats or vulnerabilities.

This research addressed several methodology stages, such as data preparation, feature engineering, model selection, model training, data splitting, and classification and evaluation. This research conducted a critical; evaluation of multiple ML algorithms, such as Decision Tree, Random Forest, and Naïve Bayes for identifying and selecting the most effective and robust ML model to attain an accurate classification of the associated attacks with DNS.

The Random Forest model emerged as the most efficient ML classification to accurately detect and classify DNS attacks with higher precision, accuracy, recall rate, and F1-score due to its ensemble nature and ability to integrate many decision trees to optimize the prediction adequacy. However, Decision Tree models also posed effective results but its slightly less effective and accurate than the Random Forest model. The Naïve Bayes model poses higher speed and simplicity but its feature-independent predictions are not always correct, particularly with the complex and higher-dimensional datasets.

This research addressed the most critical features that contributed to optimizing the accuracy of classifications and offered valuable research insights to address the key traits of DNS network traffic. Further, the correlation analysis provided a better understanding of the relation among diverse features and their influence on the predictions of ML models. The cross-validation further enhances the validity and reliability of the research outcomes and findings. The obtained research results demonstrated the chosen Random Forest ML model attained superiority in correctly classifying various DNS attacks with an accuracy of 97% and precision of 95.5%. The research findings are highly valuable for industry professionals, researchers, and practitioners in further refining the selection and utilization of ML models. This research also demonstrated the potential of integrating multiple machine learning models or utilizing a hybrid approach to further optimize the accuracy of attack detection and classification.

## Limitations

Despite the key achievements and valuable contributions of this research to enhance the detection and classification of the types of DNS attacks with the help of machine learning methods, this research also posed several limitations that might impact the generalizability and applicability of this research.

* This research mainly focused on the CIC-Bell-DNS-EXF-2021 dataset that may not comprise the entire range of the attack patterns related to domain name systems occurring across diverse network environments. Further, this dataset may not capture the entire range of complexity and diversity of the real-time DNS traffic.
* The chosen Random Forest model demonstrated higher accuracy and precision due to training on a specific dataset. Its performance largely depends on the feature selection, so it might be possible that if the features do not correctly capture the DNS traffic patterns, then its performance can be compromised (Khan & Sharma, 2023).
* The training of Random Forest might be computationally expensive and need a significant amount of memory and training that can restrict the feasibility of this ML model’s deployment.
* This research mainly emphasized the static evaluation of the ML models with the help of a pre-controlled dataset, which poses a lack of real-time testing indicating that the model performance against live atmospheres with real-world network traffic remained unvalidated.
* This research has a constrained scope, which is based on the detection of specific DNS attack types, which does not cover the many other types and forms of DNS attacks that might also impact the overall DNS applications.

## Recommendations

Based on the obtained research findings and limitations, this section provides key actionable recommendations for guiding further research and enhancing the research outcomes.

* Future research needs to focus on incorporating more nuanced and real-world DNS traffic data from different sources, including diverse types of networks, geographic locations, and user atmospheres for ultimately improving the model’s generalizability to ensure robustness against a comprehensive range of the types of attack patterns and vectors.
* Future research should emphasize validating the efficacy against real-world scenarios in live environments to get vital insights related to the model’s effectiveness in adapting to dynamic DNS traffic conditions (Peneti & E, 2021).
* Future research can employ advanced and automated feature selection methods, such as principal component analysis and recursive feature elimination for identifying the most important and relevant features to better train the ML models for eventually enhancing the accuracy of these models and reducing the issue of model overfitting.
* Future research needs to examine the methods to reduce the ensemble model’s computational complexities without affecting their accuracy. The use of a hybrid approach can largely help in balancing the accuracy of attack detection and computational proficiency.
* Future research should focus on extending this research across multiple datasets and a more extensive range of DND attacks and other network-centered attacks to realize a more efficient and robust detection framework to ultimately optimize the applicability and utility across diversified cybersecurity contexts.

## Self-Reflection

Reflecting on this research based on the classification of DNS attack types using a machine learning model on the CIC-Bell-DNS-EXF-2021 Dataset was an invaluable learning experience for me that largely improved and shaped my personal, professional, and academic growth. Undertaking this research, I have attained several learning opportunities and avenues for personal and professional growth that surely contribute to my future career. Here I outline and discuss my experience throughout this research journey and how I attained new skills and polished my existing skills and abilities.

The selection of this research topic was directed by my intense interest in the machine learning and cybersecurity research fields. I was largely encouraged by the increased emergence and prevalence of domain name systems and their susceptibility toward cyber attacks. Machine learning algorithms have immense potential to provide advanced methods for efficiently detecting and classifying cyber attacks. In the initial phase of this research, I conducted in-depth research and invested a significant amount of time to better understand this research field and set clear research aims & objectives, research methodologies, and possible challenges. This phase realized a solid foundation for further project phases. During the literature review, phase I faced several challenges due to identifying and selecting the most relevant research studies to gain an effective understanding of used theories, methodologies, and concepts in existing research. The conducted literature review largely helps in developing my analytical skills along with an ability to distinguish the strengths and weaknesses of diverse approaches. Further, it also increases my knowledge and understanding of cyber threats and the potential of machine learning methods to prevent these cyber threats. The methodology design and dataset setup were the most technically demanding sections of this whole research. I have selected Random Forest, Decision Tree, and Naïve Bayes algorithms. During this project phase, I largely improved my technical abilities, specifically in data preparation, feature engineering, model selection, and model training. I also gained proficiency in Python and its libraries, such as Pandas, Scikit, Keras, and Seaborn. I have reflected on my strong ability to deal with unforeseen challenges and manage dataset imbalances using fine-training methods. In this research, I efficiently handled the selected large and high-dimension dataset, which would definitely help in including more advanced tools and techniques. I have evaluated the performance of selected machine learning models based on accuracy, recall rate, precision, and F1 score. However, in future research, I will focus on utilizing explainable AI to realize in-depth insights through improved decision-making abilities and increase usability and trust in real-world scenarios.

I also improved my awareness of the inherent legal, ethical, and social considerations in the cybersecurity research field-based research. I have recognized the invaluable role of perseverance and resilience through the faced data limitations and technical hurdles. From this experience, I also improved my time management skills by setting realistic project goals and prioritizing the project tasks.

# References

Abad, S., Gholamy, H. and Aslani, M. (2023) ‘Classification of malicious urls using machine learning’, *Sensors*, 23(18), p. 7760. doi:10.3390/s23187760.

Alzahrani, R.J. and Alzahrani, A. (2021) ‘Security analysis of ddos attacks using machine learning algorithms in networks traffic’, *Electronics*, 10(23), p. 2919. doi:10.3390/electronics10232919.

Bakro, M. *et al.* (2024) ‘Building a cloud-ids by hybrid bio-inspired feature selection algorithms along with Random Forest Model’, *IEEE Access*, 12, pp. 8846–8874. doi:10.1109/access.2024.3353055.

Banadaki , Y.M. (2020) ‘Detecting Malicious DNS over HTTPS Traffic in Domain Name System using Machine Learning Classifiers’, *Journal of Computer Sciences and Applications.*, 8(2), pp. 46–55. doi:10.12691/jcsa-8-2-2 .

Behnke, M. *et al.* (2021) ‘Feature engineering and machine learning model comparison for malicious activity detection in the DNS-over-HTTPS protocol’, *IEEE Access*, 9, pp. 129902–129916. doi:10.1109/access.2021.3113294.

Bilge, L. *et al.* (2020) ‘EXPOSURE: Finding Malicious Domains Using Passive DNS Analysis’, *Northeastern University, Boston* [Preprint].

Bisiaux, J.-Y. (2014) ‘DNS threats and Mitigation Strategies’, *Network Security*, 2014(7), pp. 5–9. doi:10.1016/s1353-4858(14)70068-6.

Burrell, J. (2023) ‘Special issue on domain name system (DNS)’, *Journal of Cyber Policy*, 8(2), pp. 137–141. doi:10.1080/23738871.2024.2312922.

Fang, L. *et al.* (2021) ‘A comprehensive analysis of ddos attacks based on DNS’, *Journal of Physics: Conference Series*, 2024(1), p. 012027. doi:10.1088/1742-6596/2024/1/012027.

Firdaus, L. and Setiadi, T. (2023) ‘Perbandingan Algoritma naive Bayes, decision tree, Dan Knn Untuk Klasifikasi Produk Populer adidas US dengan confusion matrix’, *Jurnal Sistem Komputer dan Informatika (JSON)*, 5(2), p. 185. doi:10.30865/json.v5i2.6124.

Gonzalez Casanova, L.F. and Lin, P.-C. (2023) ‘Malicious network traffic detection for DNS over HTTPS using machine learning algorithms’, *APSIPA Transactions on Signal and Information Processing*, 12(2). doi:10.1561/116.00000058.

Gyamfi, N.K., Ceponis, D. and Goranin, N. (2022) ‘Automated system-level ANOMALY DETECTION and classification using modified Random Forest’, *2022 1st International Conference on AI in Cybersecurity (ICAIC)* [Preprint]. doi:10.1109/icaic53980.2022.9897027.

Huang, X. (2023) ‘Prediction of heart attack on Random Forest and logistic regression’, *Proceedings of the 1st International Conference on Data Analysis and Machine Learning* [Preprint]. doi:10.5220/0012816000003885.

Hubman, J.M. *et al.* (2015) ‘Ethical considerations in the cyber domain’, *Advances in Information Security*, pp. 163–174. doi:10.1007/978-3-319-23585-1\_12.

Ismail *et al.* (2022) ‘A machine learning-based classification and prediction technique for ddos attacks’, *IEEE Access*, 10, pp. 21443–21454. doi:10.1109/access.2022.3152577.

Jin, Y., Tomoishi, M. and Matsuura, S. (2019) ‘A detection method against DNS cache poisoning attacks using machine learning techniques: Work in progress’, *2019 IEEE 18th International Symposium on Network Computing and Applications (NCA)* [Preprint]. doi:10.1109/nca.2019.8935025.

Khan, A. and Sharma, I. (2023) ‘AI-enabled approach for preventing DNS attacks on Banking Institutions’, *2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE)* [Preprint]. doi:10.1109/rmkmate59243.2023.10369196.

Klein, J. *et al.* (2016) ‘A reference architecture for Big Data Systems in the National Security Domain’, *Proceedings of the 2nd International Workshop on BIG Data Software Engineering* [Preprint]. doi:10.1145/2896825.2896834.

Kozhuharova, D., Kirov, A. and Al-Shargabi, Z. (2022) ‘Ethics in cybersecurity. what are the challenges we need to be aware of and how to handle them?’, *Lecture Notes in Computer Science*, pp. 202–221. doi:10.1007/978-3-031-04036-8\_9.

Lhotka, L. and Špaček, P. (2021) *Yang types for DNS classes and resource record types* [Preprint]. doi:10.17487/rfc9108.

Liao, R. and Wang, S. (2024) ‘Malicious domain detection based on semi‐supervised learning and parameter optimization’, *IET Communications*, 18(6), pp. 386–397. doi:10.1049/cmu2.12739.

Lyu, M. *et al.* (2021) ‘Hierarchical anomaly-based detection of distributed DNS attacks on Enterprise Networks’, *IEEE Transactions on Network and Service Management*, 18(1), pp. 1031–1048. doi:10.1109/tnsm.2021.3050091.

M. Banadaki, Y. (2020) ‘Detecting malicious DNS over HTTPS traffic in domain name system using machine learning classifiers’, *Journal of Computer Sciences and Applications*, 8(2), pp. 46–55. doi:10.12691/jcsa-8-2-2.

Magalhães, V. *et al.* (2023) ‘Predicting moisture content on wood using machine learning classification methods’, *Proceedings of the 25th International Conference on Enterprise Information Systems* [Preprint]. doi:10.5220/0011988600003467.

Mahdavifar, S. *et al.* (2021) ‘Lightweight hybrid detection of data exfiltration using DNS based on machine learning’, *2021 the 11th International Conference on Communication and Network Security* [Preprint]. doi:10.1145/3507509.3507520.

Meitei, I.L., Singh, K.J. and De, T. (2016) ‘Detection of ddos DNS amplification attack using classification algorithm’, *Proceedings of the International Conference on Informatics and Analytics* [Preprint]. doi:10.1145/2980258.2980431.

MNIF GARGOURI, M. and Bouamama, S. (2024) *Intelligent transport traffic flow prediction using Machine Learning Decision tree models: An extensive analysis* [Preprint]. doi:10.2139/ssrn.4836779.

Naing, S.K. and Thwel, T.T. (2023) ‘A study of DDOS attack classification using Machine Learning Classifiers’, *2023 IEEE Conference on Computer Applications (ICCA)* [Preprint]. doi:10.1109/icca51723.2023.10182146.

Oberoi, A. and Chauhan, R. (2019) ‘Visualizing data using Matplotlib and Seaborn Libraries in Python for Data Science’, *International Journal of Scientific and Research Publications (IJSRP)*, 9(3). doi:10.29322/ijsrp.9.03.2019.p8733.

Online, S. (2023) *Top 10 powerful python libraries for data science - shiksha online*, *Study in India*. Available at: https://www.shiksha.com/online-courses/articles/top-10-powerful-python-libraries-for-data-science/#Scikit-learn (Accessed: 29 August 2024).

Ozkan-Okay, M. *et al.* (2023) ‘A novel feature selection approach to classify intrusion attacks in network communications’, *Applied Sciences*, 13(19), p. 11067. doi:10.3390/app131911067.

Panigrahi, G.R. *et al.* (2024) ‘Analytical validation and integration of CIC-Bell-DNS-EXF-2021 dataset on Security Information and Event Management’, *IEEE Access*, 12, pp. 83043–83056. doi:10.1109/access.2024.3409413.

Peneti, S. and E, H. (2021) ‘DDOS attack identification using machine learning techniques’, *2021 International Conference on Computer Communication and Informatics (ICCCI)* [Preprint]. doi:10.1109/iccci50826.2021.9402441.

PetrovaHi, B. and Petrova, B. (2024) *DNS history. when and why was DNS created?*, *ClouDNS Blog*. Available at: https://www.cloudns.net/blog/dns-history-creation-first/ (Accessed: 29 August 2024).

Pranckevičius, T. and Marcinkevičius, V. (2017) ‘Comparison of naive bayes, Random Forest, decision tree, support vector machines, and logistic regression classifiers for text reviews classification’, *Baltic Journal of Modern Computing*, 5(2). doi:10.22364/bjmc.2017.5.2.05.

Pryke, B. (2024) *How to use Jupyter Notebook in 2024: A beginner’s tutorial*, *Dataquest*. Available at: https://www.dataquest.io/blog/jupyter-notebook-tutorial/ (Accessed: 29 August 2024).

Ray, S. (2024) *Naive Bayes classifier explained: Applications and practice problems of naive Bayes classifier*, *Analytics Vidhya*. Available at: https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#:~:text=In%20simple%20terms%2C%20a%20Naive,tasks%20such%20as%20text%20classification (Accessed: 29 August 2024).

Singh, A. and Gaur, S. (2024) ‘Detection of benign and malicious DNS traffic in ISP network using machine learning algorithms’, *Recent Advances in Computational Intelligence and Cyber Security*, pp. 117–124. doi:10.1201/9781003518587-10.

Singh, R. and Srivastav, G. (2021) ‘Novel framework for anomaly detection using machine learning technique on CIC-IDS2017 dataset’, *2021 International Conference on Technological Advancements and Innovations (ICTAI)* [Preprint]. doi:10.1109/ictai53825.2021.9673238.

Song, J. (2023) ‘Comparison and analysis of accuracy of traditional random forest machine learning model and XGBoost model on Music Emotion Classification Dataset’, *2023 4th International Conference on Machine Learning and Computer Application* [Preprint]. doi:10.1145/3650215.3650340.

Stedman, C. (2024) *What is data preparation? an in-depth guide: TechTarget*, *Business Analytics*. Available at: https://www.techtarget.com/searchbusinessanalytics/definition/data-preparation (Accessed: 29 August 2024).

Thomas, D.R. *et al.* (2017) ‘Ethical issues in research using datasets of illicit origin’, *Proceedings of the 2017 Internet Measurement Conference* [Preprint]. doi:10.1145/3131365.3131389.

UNB (2022) *CIC-Bell-DNS-EXF-2021 dataset*, *University of New Brunswick est.1785*. Available at: https://www.unb.ca/cic/datasets/dns-exf-2021.html#:~:text=Furthermore%2C%20our%20proposed%20model%20could,from%20small%20to%20large%20sizes. (Accessed: 29 August 2024).

Wagan, A.A. *et al.* (2023) ‘A unified learning approach for malicious domain name detection’, *Axioms*, 12(5), p. 458. doi:10.3390/axioms12050458.

Wang, Z., Huang, J. and Rose, S. (2018) ‘Evolution and challenges of DNS-based cdns’, *Digital Communications and Networks*, 4(4), pp. 235–243. doi:10.1016/j.dcan.2017.07.005.

Wei, D. (2024) *Essential Math for Machine Learning: Confusion Matrix, accuracy, precision, recall, F1-score*, *Medium*. Available at: https://medium.com/@weidagang/demystifying-precision-and-recall-in-machine-learning-6f756a4c54ac (Accessed: 31 August 2024).

Weinand, A. *et al.* (2023) ‘Naïve Bayes supervised learning based physical layer authentication: Anti-spoofing techniques for Industrial Radio Systems’, *International Conference on Cyber Warfare and Security*, 18(1), pp. 420–430. doi:10.34190/iccws.18.1.983.

Zebin, T., Rezvy, S. and Luo, Y. (2022) ‘An explainable AI-based Intrusion Detection System for DNS over HTTPS (DOH) attacks’, *IEEE Transactions on Information Forensics and Security*, 17, pp. 2339–2349. doi:10.1109/tifs.2022.3183390.

Zhang, W. *et al.* (2016) ‘Model checking the DNS under DNS cache-poisoning attacks using spin’, *ScienceAsia*, 42S(1), p. 49. doi:10.2306/scienceasia1513-1874.2016.42s.049.

Zola, A. (2021) *What is the python programming language?*, *WhatIs*. Available at: https://www.techtarget.com/whatis/definition/Python (Accessed: 29 August 2024).

Žiža, K., Tadić, P. and Vuletić, P. (2023) ‘DNS exfiltration detection in the presence of adversarial attacks and modified exfiltrator behaviour’, *International Journal of Information Security*, 22(6), pp. 1865–1880. doi:10.1007/s10207-023-00723-w.