A Hybrid Ensemble Model using XGBoost and AdaBoost to detect and distinguish zero-day attacks

MSc Research Project
MSc Cyber Security

Ajay Krishna Edakkat Parambil
Student ID: X22110674

School of Computing
National College of Ireland

Supervisor: Vikas Sahni

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: Ajay Krishna Edakkat Parambil

Student ID: X22110674

Programme: MSc Cyber Security **Year:** 2023-2024

Module: MSc Academic Internship

Supervisor: Vikas Sahni
Submission Due Date: January 31

Project Title: A Hybrid Ensemble Model using XGBoost and AdaBoost to detect and distinguish zero-day attacks

Word Count: …………………6253………………… **Page Count** 19

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

Signature: Ajay Krishna Edakkat Parambil ……………………………………………………………………………………………………

Date: …31-01-2024………………………………………………………………………………

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

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

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

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

A Hybrid Ensemble Model using XGBoost and AdaBoost to detect and distinguish zero-day attacks

Ajay Krishna Edakkat Parambil
x22110674

Abstract

This research investigates the crucial role of Intrusion Detection Systems (IDS) in addressing cyber threats, with a specific emphasis on the detection of Zero Day assaults. Zero-day attacks, exploiting vulnerabilities concealed from developers and security experts, present a substantial security threat due to the unavailability of immediate patches. Traditionally, zero-day attack detection relied on machine learning algorithms like AdaBoost, which, while highly effective in aggregating weak learner predictions, is inefficient when dealing with complex, multi-class datasets frequently encountered in network traffic analysis.Which results in inadequate protection of critical assets, intellectual property, and sensitive data. The performance of various machine learning models is evaluated, to determine the most effective model for network intrusion detection, and to emphasise the relevance of flexibility and precision in identifying developing threats. Four unique machine learning models the AdaBoost Classifier, XGBoost Classifier, Random Forest Classifier, and a Hybrid Ensemble Model leveraging the capabilities of the above machine learning techniques for an Intrusion Detection System (IDS) that effectively identifies zero-day attacks and false positive reduction is introduced. Using a large dataset, these models are tested for their capacity to identify various network activity and, more importantly, their ability to detect Zero Day attacks. The results reveal that the Hybrid Ensemble Model achieves the highest accuracy of 82%, compared to the AdaBoost Classifier with 41%, XGBoost Classifier with 75% and Random Forest Classifier with 77%.

Keywords: Hybrid Ensemble Model, XGBoost, AdaBoost, Intrusion Detection, Zero-day attacks, machine learning

1 Introduction

In recent years, there has been a noticeable increase in cyberattacks, with zero-day attacks emerging as particularly destructive. Intrusion Detection Systems (IDS) are critical in this domain, serving as the first line of defence against a wide range of assaults. Zero-day assaults are a particularly difficult class of intrusions because of their novelty and the lack of past information or signs in training data. Detecting such assaults necessitates the use of cutting-edge tools and procedures (Zoppi et al., 2021). A zero-day attack is defined as a network traffic pattern with no corresponding patterns in network malware or attack detection components (Hindy et al., 2020). A missed attack exposes the system to immediate risk, while a false alarm can squander valuable resources by erroneously flagging benign activities as potential threats. Recent advancements in deep learning models have led to improved detection of real threats, yet they have not fully mitigated the persistent issue of false alarms, diminishing their overall efficiency. (Sun et al., 2021)

This study explores and assesses machine learning models for network intrusion detection, with a focus on detecting Zero Day attacks. Because it enables for the automatic identification of complex patterns and irregularities in network traffic data, machine learning is a possible method for boosting IDS capabilities. The performance of four machine learning models is evaluated in this paper: the AdaBoost Classifier, the XGBoost Classifier, the Random Forest Classifier, and a Hybrid Ensemble Model. The study's context encompasses the larger context of Intrusion Detection Systems (IDS) in the subject of cybersecurity. As cyber-attacks become more sophisticated, the requirement for effective intrusion detection systems (IDS) has become critical. IDSs are critical network security components because they detect and mitigate possible threats and abnormalities. It is especially critical to detect Zero Day attacks, which are distinguished by their predictability and the absence of previous indicators in training data.

An Intrusion Detection System (IDS) for identifying network anomalies and cyber threats, with a particular emphasis on detecting zero-day attacks as been developed and evaluated. The primary purpose is to assess the effectiveness of several machine learning models in recognising various network activities and attack categories, such as the AdaBoost Classifier, XGBoost Classifier, Random Forest Classifier, and a Hybrid Ensemble Model. The goal is to provide insights into the usefulness of these models in identifying both known and novel threats, ultimately contributing to network security enhancement and even the mitigation of emerging cybersecurity risks.

1.1 Research Objectives

The research objectives of this report are as follows:

1. To systematically evaluate the performance of machine learning models in the context of intrusion detection systems, such as the AdaBoost Classifier, XGBoost Classifier, Random Forest Classifier, and a Hybrid Ensemble Model.
2. To assess the model’s ability to detect Zero Day assaults, which are distinguished by their originality and absence from training data.
3. To assess the model’s classification accuracy and efficacy in distinguishing diverse network activities and attack types, a complete set of performance measures, including accuracy, precision, recall, F1-score, and confusion matrices, will be used.

1.2 Research Questions

The research questions for this report are as follows:

1. In the context of Intrusion Detection Systems (IDS), how do several machine learning models, such as the AdaBoost Classifier, XGBoost Classifier, Random Forest Classifier, and a Hybrid Ensemble Model, perform for network anomaly and cyber threat detection?
2. To what degree can these machine learning models detect and categorise Zero Day assaults, which are distinguished by their originality and lack of presence in the training data?

1.3 Research Contribution

The work reveals many research gaps in the realm of cybersecurity Intrusion Detection Systems (IDS). One major gap is the requirement for more resilient and adaptable IDS models capable of detecting Zero Day attacks, which are distinguished by their unpredictability and lack of training data. Current IDS models frequently fail to keep up with shifting attack patterns, emphasising the importance of new and robust techniques. Another research gap is the lack of investigation into model interpretability and explainability in the context of IDS. There is an increasing desire for intrusion detection systems (IDSs) to give clear and understandable insights into their decision-making processes, which can improve confidence and decision support.

2 Related Work

2.1 Hybrid ML Models in IDS

Due to the rising cybersecurity risks across multiple sectors, the area of intrusion detection systems (IDS) has seen a substantial increase in research activity in recent years. The common thread running across this research is the urgent need to build intelligent, accurate, and resilient intrusion detection systems (IDSs) to detect and neutralise growing security threats. (Sun et al., 2020) created a DL-IDS by integrating Convolutional Neural Networks (CNN) and Long Short-Term Memory Networks (LSTM) to extract spatial and temporal data for better intrusion detection.

(Al-Emadi et al., 2020) proposes a deep learning-basedintrusion detection and prevention system that uses a convolutional neural network (CNN) to extract features from network traffic and classify it as normal or malicious. Along with (Hussain et al., 2021), the use of deep learning techniques for intrusion detection showed promising results. (Kiflayet al., 2021) proposesd a NIDS that uses ensemble machine learning to improve the performance of attack detection and decrease the rate of false alarms. The performance of these presented systems was evaluated using the NSL-KDD dataset.

Similarly, (Pi et al., 2019) presented a hybrid intrusion detection system (IDS) that combines Spark ML and Convolutional-LSTM networks to handle both global and local latent threat signatures while concentrating on scalability. (Khan, 2021) suggested a Convolutional Recurrent Neural Network (CRNN) for reliable cyberattack prediction and classification, focusing on the capacity to identify both known and novel threats. (Karim et al., 2019) demonstrated a two-stage intrusion detection system with anomaly and abuse detection modules based on Spark ML and Conv-LSTM networks that achieved high accuracy. Subsequently, (Yang et al., 2021) presented a multi-tiered hybrid IDS that combines signature-based and anomaly-based techniques to identify known and undiscovered threats in vehicular networks. (Cavusoglu, 2019) also offers a hybrid intrusion detection system (IDS) with layered architectures and feature selection algorithms adapted to diverse attack types, delivering high accuracy and minimal false positives. (Ren et al., 2019) present a hybrid intrusion detection system (IDS) that incorporates Isolation Forest, genetic algorithms, and Random Forest for robust intrusion detection, solving the issue of low detection rates.

(Hassan et al., 2020) achieve outstanding results by leveraging deep learning models such as CNNs and WDLSTMs to identify network intrusions in a large data environment. (Balyan et al., 2022) created a hybrid network-based intrusion detection system that incorporates EGA-PSO and IRF algorithms for data balancing and feature selection, yielding higher performance on the NSL-KDD dataset. Finally, (Aljawarneh et al., 2018) concentrate on feature selection, using the Vote method in conjunction with several classifiers to decrease false positives and improve intrusion detection accuracy. The development of hybrid models that integrate several machine learning approaches to improve IDS performance, addressing difficulties such as data imbalance, feature selection, and the necessity for efficient intrusion detection in varied scenarios, is like these works. These methods show the power of integrating data preprocessing, and feature selection, but also classification techniques to improve intrusion detection in a variety of network contexts.

Table 2.1: Comparison Table of Hybrid ML Models in IDS

Study (Year)	Approach	Key Challenge	Main Techniques	Performance Results
Sun et al. (2020)	Deep Learning (DL)	Key challenge in this study is the unbalanced Datasets	Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM)	98.67% overall accuracy, >99.50% attack type accuracy
Pi et al. (2019)	Hybrid (Spark ML & Conv-LSTM)	Key challenge in this study is Scalability, Latent Threats	Apache Spark ML, Convolutional LSTM (Conv-LSTM)	97.29% accuracy, F1-scores 0.963 and 0.800 for zero-day
Khan (2021)	Deep Learning (CRNN)	Key challenge in this study is Unbalanced Datasets	Convolutional Recurrent Neural Network (CRNN)	97.75% accuracy, 0.963 F1-score for zero-day attacks
Karim et al. (2019)	Hybrid (Spark ML & Conv-LSTM)	Scalability, Accuracy are main key challenges	Apache Spark ML, Convolutional LSTM (Conv-LSTM)	97.29% accuracy, real-time processing <0.6 ms
Yang et al. (2021)	Hybrid (Signature-based & Anomaly-based)	Unbalanced Datasets, Real-time Processing are the main key challenges to address	Hybrid IDS approach combining signature-based and anomaly-based techniques	99.99% accuracy on CAN-intrusion-dataset, 99.88% on CICIDS2017, high F1-scores for zero-day attacks
Çavuşoğlu (2019)	Hybrid IDS with layered architecture and feature selection	Addressing different attack types	Feature selection, various machine learning algorithms	Accuracy is found to be 94% in the R2L attack type
Ren et al. (2019)	Hybrid IDS with Isolation Forest, genetic algorithms, and Random Forest	Key challenge in this study is the Low detection rates	Isolation Forest, genetic algorithms, Random Forest	92% Accuracy for AdaBoost,RUSBoost and Do_IDS

Hassan et al. (2020)	Hybrid IDS combining CNN and WDLSTM	Intrusion detection in a big data environment	Convolutional Neural Network (CNN), Weight-Dropped Long Short-Term Memory (WDLSTM)	97.15% Accuracy
Balyan et al. (2022)	Hybrid network-based IDS with EGA-PSO and IRF	Data imbalance and feature selection	Enhanced Genetic Algorithm (EGA-PSO), Improved Random Forest (IRF)	Accuracy achieved is 98.97%
Aljawarneh et al. (2018)	Hybrid approach using the Vote algorithm and various classifiers	High false positives and low false negatives	Vote algorithm, J48, Meta Pagging, RandomTree, etc.	Accuracy achieved is 99.81%

2.2 Machine Learning in IDS

To improve the performance of an AdaBoost-based IDS, (Yulianto et al., 2019) presented the Synthetic Minority Oversampling Technique (SMOTE) (Juanjuan et al., 2007), Principal Component Analysis (PCA), and Ensemble Feature Selection (EFS). (Shahraki et al., 2020) investigated the use of boosting algorithms for intrusion detection systems, especially Real Adaboost, Gentle Adaboost, and Modest Adaboost, with an emphasis on performance and stability. In a similar line, (Rachmadi et al., 2021) improved DoS detection in IoT systems using AI and the AdaBoost algorithm, reaching excellent accuracy and precision. These research all had the same goal: to improve IDSs using advanced methodologies, whether by addressing data imbalance, improving algorithm selection, or using AI. The use of machine learning methods, such as AdaBoost, to improve the performance of IDSs and the necessity to address difficulties such as data imbalance and attack detection accuracy are commonalities among these approaches.

(Devan et al., 2020) and (Bhattacharya et al., 2020) argue for the use of machine learning models in network intrusion categorization, highlighting the need of accuracy and precision. (Alzahrani et al., 2021) apply machine learning techniques for intrusion detection as well, with an emphasis on feature engineering and data pretreatment. (Dang, 2019) proposes tree-based ensemble learning as an effective strategy for IDS, whereas (Dhaliwal et al., 2018) strive to comprehend and analyse network data in order to construct robust Intrusion Detection Systems. (Faysal et al., 2022) offer a hybrid machine learning approach for IoT security, focusing on the detection of IoT device threats. As a result, these research all have the same objective of applying machine learning to improve network and IoT security, with an emphasis on attaining accuracy, precision, and efficacy in intrusion detection, addressing the growing danger of network assaults in a dynamic technological context.

Table 2.1: Comparison Table of ML Models in IDS

Study (Year)	Approach	Key Challenge	Main Techniques	Performance Results
Yulianto et al. (2019)	Enhancing AdaBoost-based IDS	Data imbalance and inappropriate classification	SMOTE, PCA, EFS	AUROC of 92% for PCA and SMOTE, precision achieved 81.83%, accuracy achieved 81.83%, recall achieved 100%, and F1 Score achieved 90.01% for EFS and SMOTE
Shahraki et al. (2020)	Evaluating Boosting Algorithms for IDS	Algorithm selection and performance stability	Real Adaboost, Gentle Adaboost, Modest Adaboost	Real and Gentle AdaBoost with approximately 70% lower error rates which is actually compared to Modest AdaBoost, Modest AdaBoost being approximately 7% faster
Rachmadi et al. (2021)	AI-Based IDS for IoT DoS Detection	DoS attacks in IoT systems	AdaBoost with AI	DoS detection accuracy of 95.84%.
Devan et al. (2020)	XGBoost-DNN model for intrusion detection	Security in the era of data proliferation	XGBoost, DNN, feature selection, normalization	97.1% Accuracy
Bhattacharya et al. (2020)	PCA-Firefly-based model for IDS	Data collection and efficient classification	PCA, Firefly algorithm, XGBoost, dataset preprocessing	99.9% Accuracy
Alzahrani et al. (2021)	Machine learning for NIDS in SDN	Data scarcity and data breach prevention	Decision Tree, Random Forest, XGBoost, preprocessing	95.5% Accuracy
Dang (2019)	Ensemble learning for IDS	Adequate labeled data collection	Tree-based ensemble learning, feature engineering	98.7% Accuracy

Dhaliwal et al. (2018)	Data analysis for IDS in SDN	Data security and privacy in network transitions	Network data analysis, improving IDS	98.7% Accuracy
Faysal et al. (2022)	XGB-RF hybrid scheme for IoT security	Resource constraints and attack detection in IoT	XGBoost, Random Forest, feature selection	99.94% detection of attacks, outperformed other methods

3 Research Methodology

3.1 Research Steps

A literature review of related and similar works were done to find out the prior works which were carried out to mitigate the issue. And similar works which were done using parts of the same methodologies and different methodologies were checked in rigorous methods to extract whatever was usefull and relevant to the present work was found. These works helped in the overall assessment of various technologies and their advantages and drawbacks compared to each other in various conditions.

3.2 Equipments and Tools Used

The work is carried out in Python3 using Google Colaboratory since it was found to be more feasible and easier to work while using multiplatforms. To help the machine learning and data analysis operations, many libraries and modules have been loaded. Among these libraries are the following: NumPy and Pandas for data manipulation, Seaborn for data visualisation, imbalanced-learn (imblearn) for addressing class imbalance using SMOTE, scikit-learn for various machine learning functionalities, matplotlib for plotting and charting, XGBoost for gradient boosting, mlxtend for stacking classifier implementation, and other specific modules for tasks such as classification reports, train-test splitting, decision tree and random forest classifiers label.

Package	Version
numpy	1.23.5
pandas	1.5.3
seaborn	0.12.2
matplotlib	3.7.1
xgboost	2.0.2
mlxtend	0.22.0
seaborn	0.12.2

3.3 Data Collection

The dataset is from Kaggle[1], UNSW-NB15 a network intrusion dataset that contains raw network packets. It was created by the IXIA PerfectStorm tool in the Cyber Range Lab of the Australian Centre for Cyber Security at the University of New South Wales (UNSW) Canberra for generating a hybrid of real modern normal activities and synthetic contemporary attack behaviors. A data training set named "UNSW_NB15_training-set.csv" was taken from

[1] https://www.kaggle.com/datasets/mrwellsdavid/unsw-nb15/

this dataset. A full set of 49 features with matching class labels was extracted from the raw network packet data of the UNSW-NB15 dataset in the feature extraction part of this project. These characteristics were created with the use of twelve different algorithms and tools, including Argus and Bro-IDS, which provide insights into various aspects of network behaviour. The retrieved features included critical criteria for network traffic analysis, allowing network activity to be classified into nine separate attack types, including Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, and Worms. The goal of the feature extraction procedure was to collect and depict network behaviours, trends, and anomalies that would be useful for future machine learning and cybersecurity study. These characteristics were saved in the dataset, laying the groundwork for developing predictive models and undertaking in-depth network security research.

3.4 Data-Pre Processing

The raw UNSW-NB15 dataset was subjected to numerous critical data preparation stages in the data preprocessing phase to make it acceptable for machine learning and analysis. The tasks included data cleansing, addressing missing information, and deleting duplicates. To ensure algorithm compatibility, categorical data, including attack kinds, was frequently encoded or translated into numerical representations. Furthermore, data normalisation or scaling was undertaken to verify that the scales of the features were similar for model training. Class imbalance concerns in cybersecurity datasets were addressed using approaches such as resampling or oversampling with methods such as SMOTE to establish a balanced distribution of attack and normal instances. To ease model evaluation, the dataset was divided into two parts: training and testing.

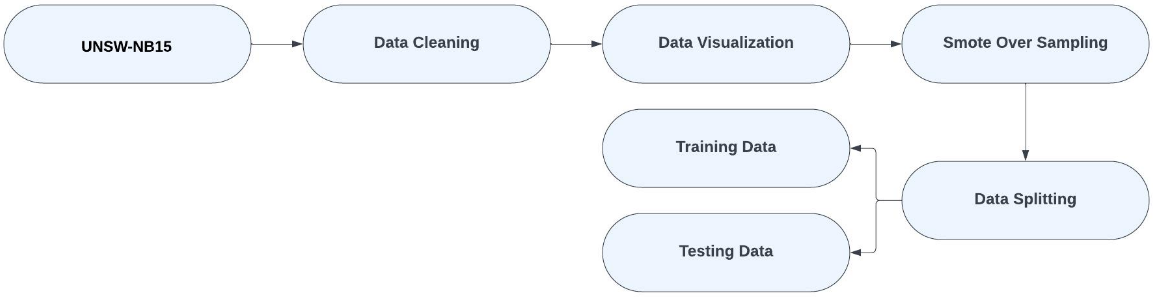

Figure 3.1: Data Pre Processing

4 Design Specification

4.1 Dataset Description

For developing an Intrusion Detection System (IDS) utilising the UNSW-NB15 dataset. The project requires the usage of specified software components, specifically Python versions less than 3.6.3 and key Python libraries such as Pandas, NumPy, XGBoost, scikit-learn, Matplotlib, Seaborn, and mlxtend, among others. The dataset, obtained from Kaggle, consists of raw network packets created by the IXIA PerfectStorm programme, which include a mix of real-world network activity and simulated attack behaviours. It consists of nine separate attack types and 49 characteristics produced from the use of twelve algorithms, each with its own class label. The dataset is divided into four CSV files: UNSW-NB15 1.csv, UNSW-NB15 2.csv, UNSW-NB15 3.csv, and UNSW-NB15 4.csv, along with ground truth and event files. Furthermore, the dataset must be segmented into a training set of 1,75,341 records and a testing set of 82,332 records in order to create and evaluate machine learning models.

5 Implementation

A variety of machine learning models was used to increase the prediction capabilities of Intrusion Detection Systems (IDS) aimed at identifying Zero Day assaults in the model training portion. As important classifiers, the AdaBoost Classifier, XGB Classifier, and Random Forest Classifier were utilised, each designed to capture different aspects of network traffic behaviour and irregularities. A Hybrid Ensemble Model, which integrated the characteristics of Random Forest , Decision Tree classifiers and AdaBoost as base models with an XGB Classifier meta-classifier, was also created. This hybrid ensemble approach aimed to improve the overall robustness and accuracy of the IDS system by pooling predictions from diverse models.

The practical implementation of the Intrusion Detection System (IDS) with the UNSW-NB15 dataset. It starts with setting up the development environment, which includes generating a Python environment with the required libraries and ensuring compatibility with the specified software versions. This stage also includes installing any extra dependencies and programmes necessary for data analysis and the development of machine learning models. The project then moves on to data preparation, which involves data cleansing, missing value management, and categorical variable encoding to preparing the dataset for machine learning.

The UNSW-NB15 dataset was separated into two distinct subsets, a training set and a testing set, using an 80-20 split ratio, with 80 percent of the data allotted to the training set and the remaining 20 percent allocated to the testing set. The ensuing machine learning model construction and validation processes relied heavily on this data separation. The training set, which contained 80% of the records, was used to train and create prediction models, allowing them to discover patterns and linkages in the data. Meanwhile, the testing set, which comprised 20% of the records, acted as an independent dataset for model evaluation and validation, allowing evaluation of the model's performance on unseen data to judge its generalizability and efficacy in real-world circumstances.

Additionally, the Synthetic Minority Oversampling Technique (SMOTE) is used to balance the distribution of attack subcategories in the dataset. The creation of machine learning models is at the heart of the implementation. The AdaBoost Classifier, XGBoost Classifier, Random Forest Classifier, and a Hybrid Ensemble Model are the four basic models used. The Hybrid Ensemble Model combines the strengths of the Random Forest and Decision Tree base models, as well as XGBoost as the meta-classifier, to improve the system's overall prediction capabilities. These models are trained using the preprocessed training dataset and features derived from the UNSW-NB15 dataset. By learning from previous network traffic data, model training tries to equip the IDS with the capacity to recognise all sorts of assaults, including Zero Day attacks. Following that, performance indicators like Accuracy Score, Confusion Matrix, Classification Report, Specificity, and Sensitivity are used to assess the models. These metrics give insight into the model’s capacity to detect and categorise various attack types, with a particular emphasis on their success in detecting Zero Day assaults, which are intrinsically difficult to detect due to their previously undetected nature. Models may be tuned and optimised to increase their accuracy and generalizability. The depiction of significant insights and results throughout the deployment process allows for a better understanding of the dataset and the model’s behaviour. Visualization approaches based on libraries such as Matplotlib and Seaborn assist in expressing and interpreting findings.

5.1 List of Models

Random Forest Classifier: The Random Forest Classifier is a strong ensemble learning approach that makes predictions by combining many decision trees. It is well-known for its durability and capacity to handle complicated and high-dimensional data. It can record a wide range of network activity patterns in the context of IDS, making it helpful in spotting both known and potentially new assaults.

Decision Tree Classifier: The Decision Tree Classifier is significantly used in the Hybrid Ensemble model. Decision trees are simple yet intuitive models that separate data based on attributes to make predictions. When employed in an IDS environment, they may identify certain patterns in network traffic data.

XGBoost Classifier: XGBoost is a versatile and high-performance gradient boosting algorithm. It is the meta-classifier in the Hybrid Ensemble paradigm. XGBoost can increase the overall forecast accuracy of the IDS system by learning from the outputs of other classifiers and making more knowledgeable final predictions.

Hybrid Ensemble Model: The Hybrid Ensemble Model is composed of a number of base classifiers, including Random Forest, Decision Tree and AdaBoost as well as a meta-classifier called XGBoost. This ensemble technique leverages the capabilities including both base classifiers and the meta-classifier to improve overall model performance. The basic classifiers identify certain peculiarities in network traffic, while the meta-classifier combines their predictions, improving the system's detection of Zero Day assaults.

5.2 Data Visualization

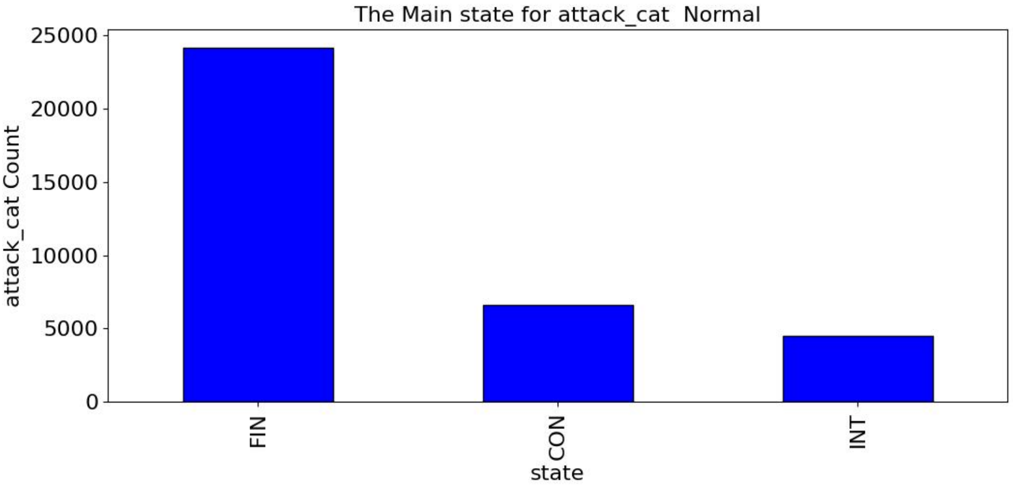

Figure 5.1: Distribution of Top 3 Network States (FIN, CON, INT) for 'Normal' Attack Category

Figure 5.1 shows a bar graph with the state-wise distribution of the top three most frequent values (FIN, CON, INT) on the x-axis and the count of attack category occurrences (attack cat Count) ranging from 0 to 25,000 on the y-axis. This graph focuses on the "Normal" attack category and depicts the distribution of the major states connected with it.

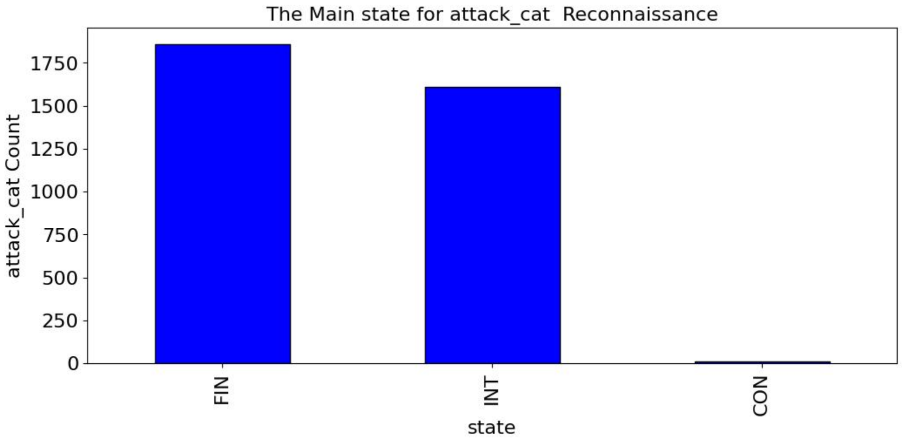

Figure 5.2: Distribution of Top 3 Network States (FIN, INT, CON) for 'Reconnaissance' Attack Category

Figure 5.2 shows a bar graph with the state-wise distribution of the top three most frequent values (FIN, INT, CON) on the x-axis and the count of occurrences in the "Reconnaissance" attack category (attack cat Count) ranging from 0 to 1750 on the y-axis. This graph focuses on the attack category "Reconnaissance," offering information on the distribution of main network states connected with this category. The heights of the bars represent counts, demonstrating how often the selected states are in the "Reconnaissance" category.

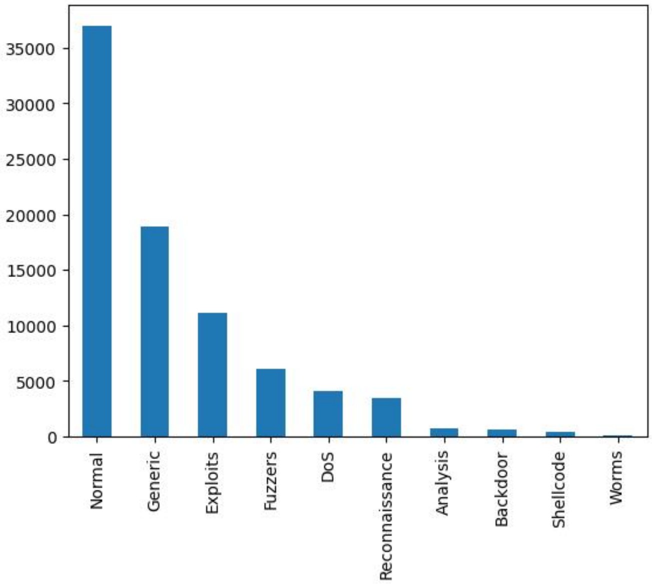

Figure 5.3: Distribution of Target Class Values

The chart in Figure 5.3 depicts the distribution of distinct classes in the dataset by providing a visual representation of the value counts within the target class. It is clear that the dataset has a class imbalance, with certain target classes greatly outnumbering others. This imbalance might complicate machine learning model training since the model may become biassed toward the majority class, resulting in inferior detection of minority classes, such as certain types of cyberattacks. To address this issue and improve the model's ability to detect all types of attacks effectively, the SMOTE method can be employed. SMOTE helps create synthetic instances of the minority class by interpolating between existing data points, thereby balancing the class distribution and mitigating the bias.

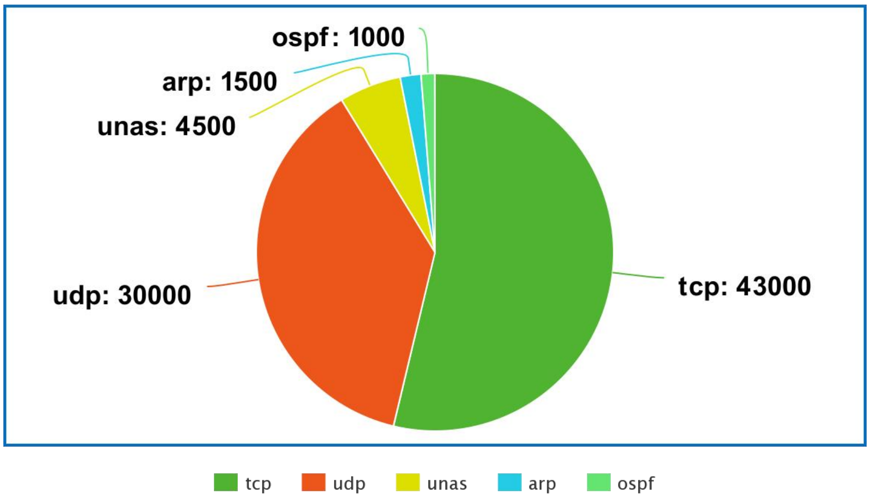

Figure 5.4: Attack Frequency

In Figure 5.4, a pie chart is presented, representing various network protocols, including TCP, UDP, UNAS, ARP, and OSPF and the frequency or count of occurrences associated with each protocol. This Line chart offers a visual representation of the distribution and prevalence of these network protocols within the dataset.

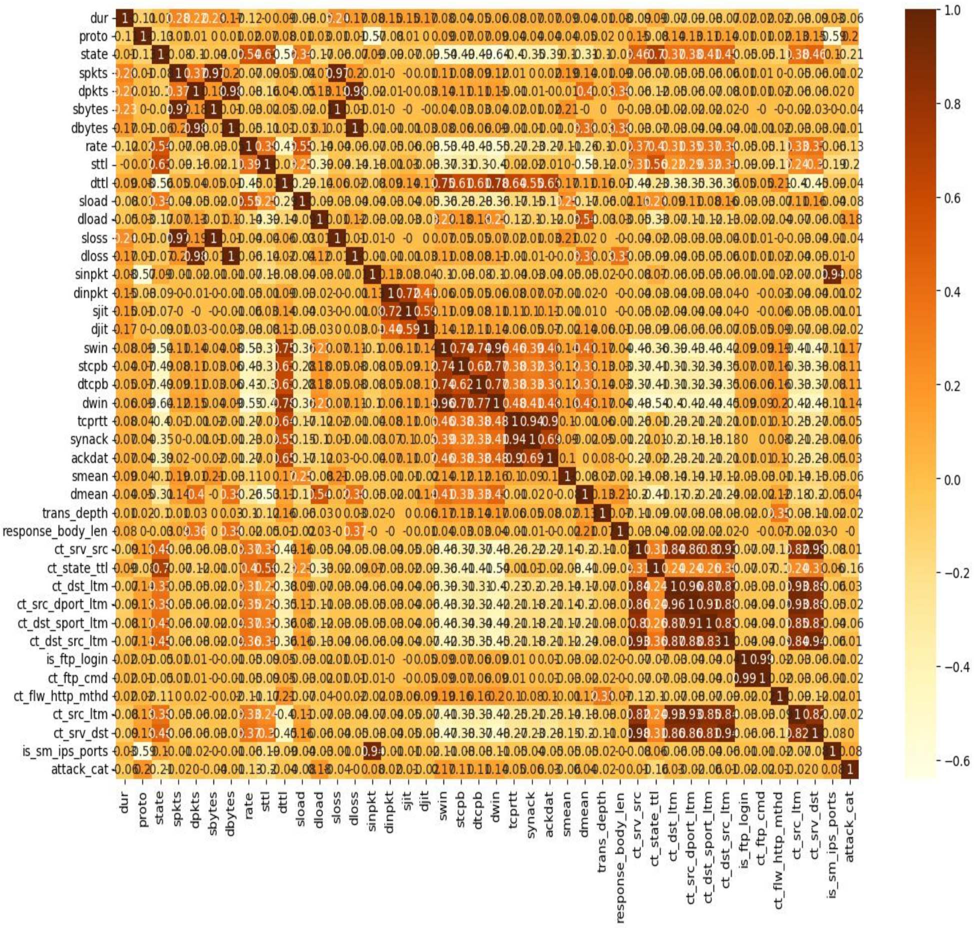

Figure 5.5 Correlation Matrix

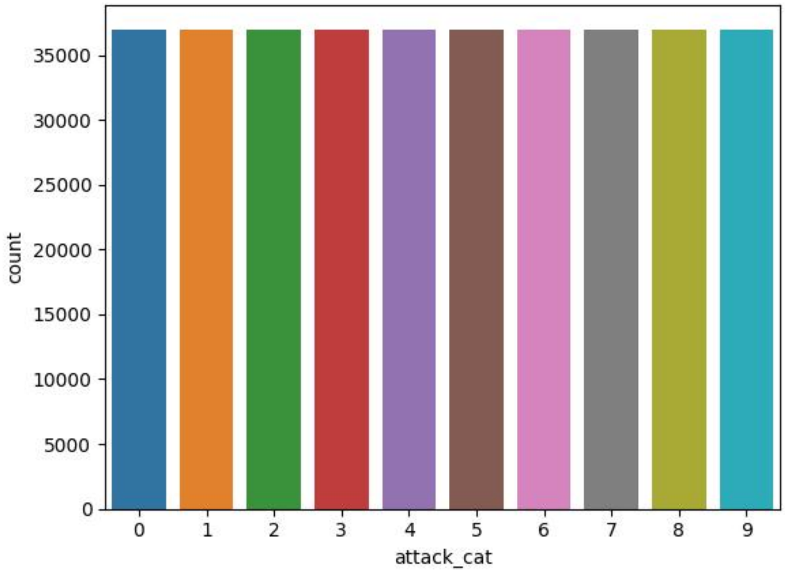

Figure 5.6: Data Balancing Using SMOTE Over-Sampling Technique

Figure 5.6 conveys the content and purpose of the visualization, indicating that it showcases the data balancing process using the SMOTE over-sampling technique. This technique is commonly employed to address class imbalance in datasets by generating synthetic instances of the minority class, thereby creating a more balanced distribution.

6 Evaluation

6.1 AdaBoost Classifier Model

AdaBoost is an ensemble learning approach that combines the outputs of numerous weak classifiers, often decision trees, to generate a powerful, adaptive classifier. Its importance in the IDS stems from its capacity to detect innovative and previously unknown attack patterns, such as Zero Day assaults, by repeatedly re-weighting and merging the classifiers to favour the right classification of misclassified cases. AdaBoost's versatility makes it ideal for rapidly changing cyber threat scenarios. By incorporating AdaBoost alongside other classifiers in a Hybrid ML model, it contributes to the system's resilience against unknown threats, improving overall detection capability by learning and adapting to ever-changing attack behaviours, and ultimately strengthening the IDS's defence mechanisms and capacity to uncover and mitigate Zero Day attacks.

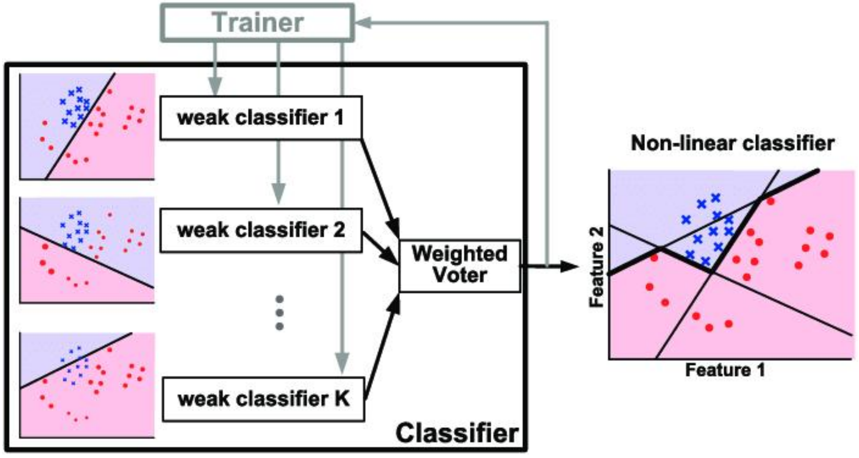

Figure 6.1: AdaBoost Classifier Architecture

The AdaBoost Classifier had an accuracy score of about 0.4061, which corresponds to about 40.61 percent. This score shows that the model accurately categorised about 40.61% of the cases in the dataset. In the context of classification tasks, accuracy is defined as the proportion of properly identified cases out of a total number of occurrences. An accuracy score of 40.61 percent, on the other hand, indicates that the model's performance may be restricted, and it may not be extremely successful in accurately categorising the data, underlining the significance of future model review and potential improvements. (Wang, Zhang and Verma, 2015)

6.2 XGB Classifier Model

XGBoost is a sophisticated gradient boosting technique that is well-known for its high performance and versatility. The XGBoost Classifier acts as the meta-classifier in the IDS architecture, making final judgments based on the outputs of other basic classifiers. This ensemble technique improves the system's robustness and forecast accuracy, enabling it to identify Zero Day assaults, which are distinguished by their unpredictability and uniqueness. XGBoost excels in capturing subtle correlations and patterns in data, and its versatility makes it especially well-suited to detecting new and previously unknown attack activities.

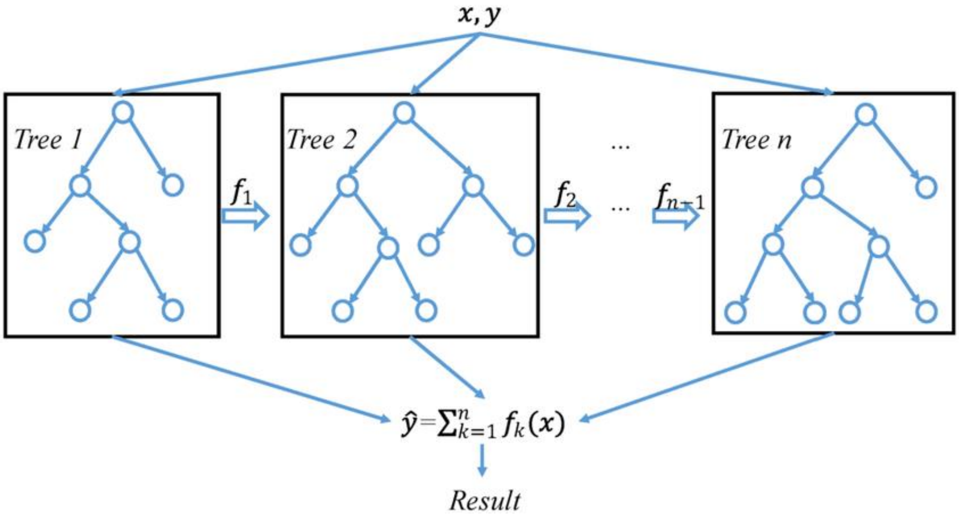

Figure 6.2: XGB Classifier Architecture

The XGBoost (Extreme Gradient Boosting) Classifier had an accuracy score of about 0.7525, which corresponds to about 75.25 percent. This score shows that the model accurately categorised about 75.25% of the cases in the dataset. In the context of classification tasks, accuracy is defined as the proportion of properly identified cases out of a total number of occurrences. A higher accuracy score, such as 75.25 percent, indicates that the XGBoost Classifier performed rather well in accurately categorising the data. (Wang et al., 2019)

6.3 Random Forest Classifier Model

Random Forest is a powerful ensemble learning approach that mixes many decision trees to produce a robust and effective classifier. Random Forest is an IDS framework component that captures and interprets detailed patterns in network traffic data, which is critical for spotting both known and novel Zero Day attacks. Its ability to handle high-dimensional and complicated data, as well as its ability to avoid overfitting, makes it an excellent candidate for the IDS's machine learning component. Random Forest works with other basic classifiers, such as Decision Trees, as part of the Hybrid ML model to capture certain subtleties in network activity, whereas this meta-classifier, XGBoost, refines the decision-making process.

The Random Forest Classifier has an accuracy score of about 0.7706, which corresponds to about 77.06 percent. This accuracy score shows that the model categorised about 77.06 percent of the cases in the dataset correctly. In the context of classification tasks, accuracy is defined as the proportion of properly identified cases out of a total number of occurrences. An accuracy score of 77.06 percent indicates that the Random Forest Classifier performed well in accurately categorising the data, displaying a high level of accuracy in its predictions. (Le et al., 2021)

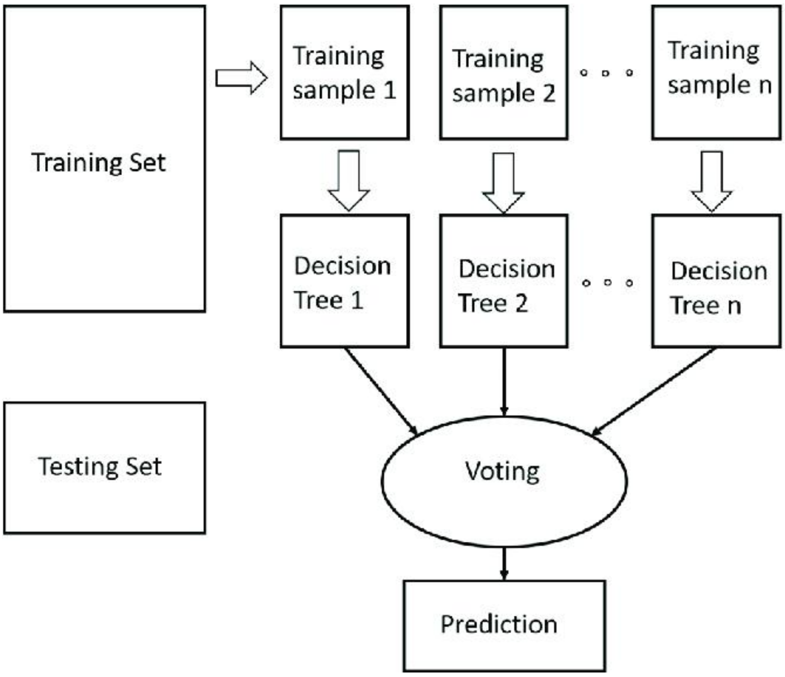

Figure 6.3: Random Forest Classifier Architecture

6.4 Hybrid Ensemble Model

The Hybrid Ensemble Model, which combines Random Forest,Decision Tree classifiers and Adaboost as base models with the meta-classifier XGBoost, provides a comprehensive and resilient threat detection technique. The fundamental classifiers, AdaBoost, Random Forest and Decision Trees, collaborate to capture subtle patterns in network data, allowing the system to recognise both known and previously unknown attack behaviours, which is important for detecting Zero Day attacks. These foundation models add to the ensemble's variety by capturing certain peculiarities in network data. XGBoost, the meta-classifier, then refines and synthesises the predictions from the underlying models, improving the IDS's overall predicted accuracy and flexibility.

The accuracy score of the Hybrid Ensemble Classifier was roughly 0.8192, which corresponds to about 81.92 percent. This accuracy score shows that the model categorised about 81.92 percent of the cases in the dataset correctly. In the context of classification tasks, accuracy is defined as the proportion of properly identified cases out of a total number of occurrences. An accuracy score of 81.92 percent indicates that the Hybrid Ensemble Classifier performed well in accurately identifying the data.

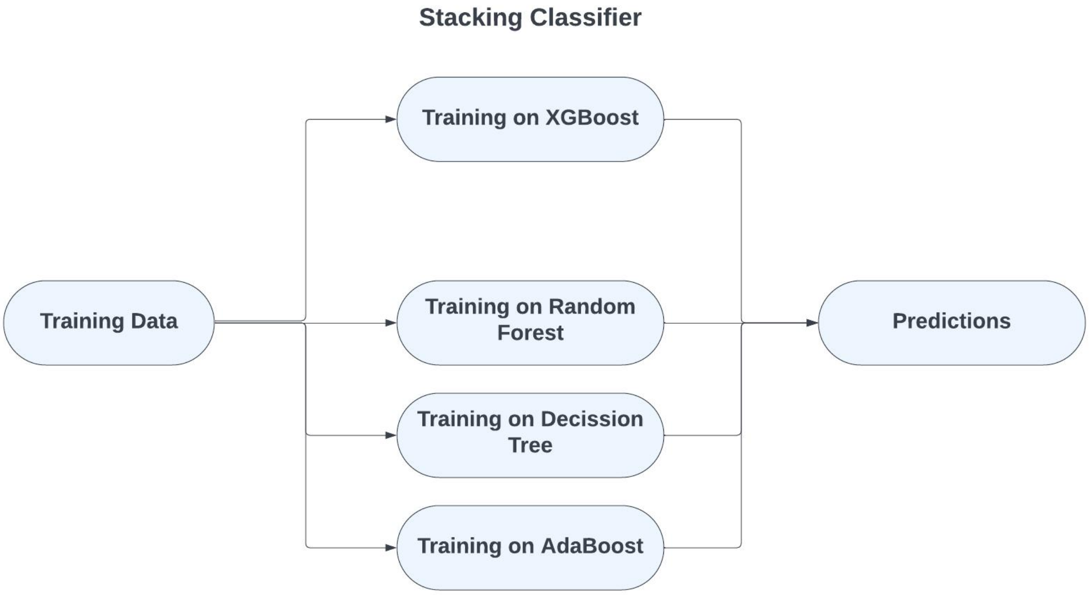

Figure 6.4: Hybrid Ensemble Architecture

6.5 Classification Performance of Machine Learning Models

In the context of Intrusion Detection Systems (IDS), the classification performance of machine learning models is evaluated using important metrics such as precision, recall, and F1-score, as well as overall accuracy. The AdaBoost Classifier has a balanced but lower accuracy of about 40.61 per cent, highlighting its possible limits of inappropriately categorising cases across many categories. In comparison, the XGBoost Classifier has a significantly greater accuracy of roughly 75.25 per cent, indicating its strength in making accurate predictions. The Random Forest Classifier comes in second with an accuracy of about 77.06 per cent, demonstrating its effectiveness in classifying attacks. The Hybrid Ensemble Model, on the other hand, exceeds the others with an accuracy of about 81.92 per cent, highlighting its resilience in generating exact classifications.

.

Model	Precision	Recall	f1-Score	Accuracy Score
AdaBoost Classifier	.48	.42	.36	0.4061
XGBoost Classifier	.78	.75	.74	0.7525
Random Forest Classifier	.79	.77	.76	0.7706
Hybrid Ensemble Model	.83	.82	.82	0.8192

Table 6.1: Comparison of Machine Learning Model Accuracy Scores for Intrusion Detection Systems

7 Conclusion and Future Work

The research looks into the world of Intrusion Identification Systems (IDS) for the detection of network abnormalities and cyber threats, with a particular emphasis on recognising zero-day assaults. The assessment and comparison of machine learning models revealed useful information about their categorization performance. The Hybrid Ensemble Model stood out among the models, getting the highest accuracy score of 82 per cent. This model combines the power of Random Forest, Decision Tree and AdaBoost classifiers with the meta-classifier XGBoost, resulting in outstanding robustness and precision in identifying a wide range of network traffic and assaults. However, it is critical to recognise the study's shortcomings. The model's performance metrics fluctuated between attack categories, emphasising the difficulty of establishing consistent accuracy in a multi-class categorization scenario. Although comprehensive, the dataset may not completely represent the expanding environment of cyber threats, since IDSs must constantly adapt to new attack patterns.

In the future, the emphasis should be on improving model interpretability and explainability, as these qualities are critical in the context of cybersecurity. Using sophisticated approaches such as feature significance analysis and model-agnostic interpretability tools can assist in understanding the model's decision-making processes. Incorporating real-time data streams and utilising cutting-edge anomaly detection algorithms can also improve an IDS's capacity to respond to emerging threats. IDSs are still on their quest to properly identify and mitigate cyber threats. The Hybrid Ensemble Model offers a viable basis for further study, exploration, and improvement, with the ultimate objective of improving network security and protecting against more complex assaults.

References

Zoppi, T., Ceccarelli, A. and Bondavalli, A. (2021) 'Unsupervised algorithms to detect zero-day attacks: Strategy and application', *IEEE Access*, 9, pp.90603-90615.

Hindy, H., Atkinson, R., Tachtatzis, C., Colin, J.N., Bayne, E. and Bellekens, X.(2020) 'Utilising deep learning techniques for effective zero-day attack detection', *Electronics*, 9(10), p.1684.

Sun, Y., Song, C., Yu, S., Pan, H., Li, T. and Liu, Y. (2021) 'A Novel Genetic Algorithm-XGBoost Based Intrusion Detection Method', *IMCEC 2021 - IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference,* pp. 51–57.

Sun, P., Liu, P., Li, Q., Liu, C., Lu, X., Hao, R. and Chen, J.(2020) 'DL-IDS: Extracting features using CNN-LSTM hybrid network for intrusion detection system', *Security and communication networks,* 2020, pp.1-11.

Al-Emadi, S., Al-Mohannadi, A. and Al-Senaid, F. (2020) 'Using Deep Learning Techniques for Network Intrusion Detection', *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies, ICIoT 2020*, pp. 171–176.

Hussain, J. and Hnamte, V. (2021) 'Deep Learning Based Intrusion DetectionSystem:Modern Approach', *2021 2nd Global Conference for Advancement in Technology, GCAT 2021 [Preprint]*.

Kiflay, A.Z., Tsokanos, A. and Kirner, R. (2021) ‘A network intrusion detectionsystemusing ensemble machine learning’, *Proceedings - International Carnahan Conference on Security Technology,* 2021-October.

Khan, I.A., Pi, D., Khan, Z.U., Hussain, Y. and Nawaz, A. (2019) ‘HML-IDS: A hybrid-multilevel anomaly prediction approach for intrusion detection in SCADA systems’, *IEEE Access,* 7, pp.89507-89521.

Khan, M.A. (2021) ‘HCRNNIDS: Hybrid convolutional recurrent neural network-based network intrusion detection system’, *Processes*, 9(5), p.834.

Khan, M.A., Karim, M.R. and Kim, Y. (2019) ‘A scalable and hybrid intrusion detection system based on the convolutional-LSTM network’, *Symmetry*, 11(4), p.583.

Yang, L., Moubayed, A. and Shami, A. (2021) ‘MTH-IDS: A multitiered hybrid intrusion detection system for internet of vehicles’, *IEEE Internet of Things Journal*, 9(1), pp.616-632.

Çavuşoğlu, Ü. (2019) ‘A new hybrid approach for intrusion detection using machine learning methods’, *Applied Intelligence*, 49, pp.2735-2761.

Ren, J., Guo, J., Qian, W., Yuan, H., Hao, X. and Jingjing, H. (2019) ‘Building an effective intrusion detection system by using hybrid data optimization based on machine learning algorithms’, *Security and communication networks*, 2019.

Hassan, M.M., Gumaei, A., Alsanad, A., Alrubaian, M. and Fortino, G. (2020) ‘A hybrid deep learning model for efficient intrusion detection in big data environment’, *Information Sciences*, 513, pp.386-396.

Balyan, A.K., Ahuja, S., Lilhore, U.K., Sharma, S.K., Manoharan, P., Algarni, A.D., Elmannai, H. and Raahemifar, K. (2022) ‘A hybrid intrusion detection model using ega-pso and improved random forest method’, *Sensors*, 22(16), p.5986.

Aljawarneh, S., Aldwairi, M. and Yassein, M.B. (2018) ‘Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model’, *Journal of Computational Science*, 25, pp.152-160.

Yulianto, A., Sukarno, P. and Suwastika, N.A, (2019) ‘Improving adaboost-based intrusion detection system (IDS) performance on CIC IDS 2017 dataset’, *In Journal of Physics: Conference Series* (Vol. 1192, p. 012018). IOP Publishing.

Juanjuan, W., Mantao, X., Hui, W. and Jiwu, Z. (2007) ‘Classification of imbalanced data by using the SMOTE algorithm and locally linear embedding’, *In International Conference on Signal Processing Proceedings, ICSP.*

Shahraki, A., Abbasi, M. and Haugen, Ø. (2020) ‘Boosting algorithms for network intrusion detection: A comparative evaluation of Real AdaBoost, Gentle AdaBoost and Modest AdaBoost’, *Engineering Applications of Artificial Intelligence*, 94, p.103770.

Rachmadi, S., Mandala, S. and Oktaria, D. (2021) ‘Detection of DoS attack using AdaBoost algorithm on IoT system’, *In 2021 International Conference on Data Science and Its Applications (ICoDSA)* (pp. 28-33). IEEE.

Devan, P. and Khare, N. (2020) ‘An efficient XGBoost–DNN-based classification model for network intrusion detection system’, *Neural Computing and Applications,* 32, pp.12499-12514.

Bhattacharya, S., Maddikunta, P.K.R., Kaluri, R., Singh, S., Gadekallu, T.R., Alazab, M. and Tariq, U. (2020) ‘A novel PCA-firefly based XGBoost classification model for intrusion detection in networks using GPU’, *Electronics*, 9(2), p.219.

Alzahrani, A.O. and Alenazi, M.J. (2021) ‘Designing a network intrusion detection system based on machine learning for software defined networks’, *Future Internet*, 13(5), p.111.

Dang, Q.V. (2019) ‘Studying machine learning techniques for intrusion detection systems’, *In Future Data and Security Engineering: 6th International Conference, FDSE 2019, Nha Trang City, Vietnam, November 27–29, 2019*, Proceedings 6 (pp. 411-426).

Dhaliwal, S.S., Nahid, A.A. and Abbas, R. (2018) ‘Effective intrusion detection system using XGBoost. Information’, 9(7), p.149.

Faysal, J.A., Mostafa, S.T., Tamanna, J.S., Mumenin, K.M., Arifin, M.M., Awal, M.A., Shome, A. and Mostafa, S.S. (2022, January) ‘XGB-RF: A hybrid machine learning approach for IoT intrusion detection’, *In Telecom* (Vol. 3, No. 1, pp. 52-69). MDPI.

Wang, Z., Zhang, J. and Verma, N. (2015) ‘Realizing Low-Energy Classification Systems by Implementing Matrix Multiplication Directly Within an ADC’, *IEEE Transactions on Biomedical Circuits and Systems*, pp. 1–1.

Wang, Y., Pan, Z., Zheng, J., Qian, L. and Li, M. (2019) ‘A hybrid ensemble method for pulsar candidate classification’, *Astrophysics and Space Science*, 364(8), p. 139.

Le, T.M., Vo, T.M., Pham, T.N. and Dao, S.V.T. (2021) ‘A Novel Wrapper–Based Feature Selection for Early Diabetes Prediction Enhanced With a Metaheuristic’, *IEEE Access*, 9, pp. 7869–7884.