Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 4, April 2024

Classification and Prediction Technique for DDoS Attacks Using Machine Learning

Meghana Lokhande
Computer Department
Pimpri Chinchwad College of Engineering, Pimpri, India

Harsh Dandge
Computer Department
Pimpri Chinchwad College of Engineering, Pimpri, India

Viraj Jadhao
Computer Department
Pimpri Chinchwad College of Engineering, Pimpri, India

Swapnil Patil
Computer Department
Pimpri Chinchwad College of Engineering, Pimpri, India

Sarvesh Powar
Computer Department
Pimpri Chinchwad College of Engineering, Pimpri, India

***Abstract*— The paper examines the use of Machine LearningML algorithms for classifying and predicting distributed denials of service attacks. DDoS attacks continue to pose significant threats to network security, making timely detection and mitigation crucial. ML algorithms offer promising capabilities in identifying and predicting such attacks. This survey paper provides a comparative analysis of popular ML algorithms, including XGBoost, RandomForest, and Naive Bayes, in terms of their effectiveness in DDoS attack detection. Additionally, a proposed method utilizing RandomForest is presented, along with a comprehensive evaluation of its performance. The study incorporates numerical data analysis and relevant diagrams to offer insights into the comparative efficacy of different ML techniques for DDoS attack detection.**

***Keywords*— DDoS attacks, machine learning, random forest, XGBoost.**

I. INTRODUCTION

Distributed Denial of Service (DDoS) attacks aim to disrupt the normal functioning of a network or service by overwhelming it with a flood of malicious traffic. Traditional defense mechanisms are often inadequate in mitigating DDoS attacks due to their evolving nature and scale. Therefore, there is a growing interest in leveraging machine learning (ML) techniques for the early detection and prediction of DDoS attacks. This paper aims to provide a comprehensive review of ML-based classification and prediction techniques for DDoS attacks, focusing on the comparative analysis of XGBoost, RandomForest, and Naive Bayes algorithms.

DDoS attacks are a growing concern for network security. These attacks involve overwhelming a network with traffic, making it unavailable to legitimate users. Traditional security measures, such as firewalls and intrusion detection systems, are often ineffective against DDoS attacks. Machine Learning (ML) techniques have been proposed as a potential solution to this problem. ML algorithms can learn patterns in network traffic and identify the possibility of a DDoS attack. In this study, we investigate the application of machine learning approaches to classify and forecast DDoS attacks. We present a comparative study of XGBoost, RandomForest, and Naive Bayes algorithms, highlighting their strengths and weaknesses in detecting DDoS attacks. We also propose a method using RandomForest for DDoS attack detection and prediction. Our method is evaluated using numerical date and Scopus index, to support our findings.

Fig. 4.1: Various types of DDoS Attacks

Distributed denial of service (DDoS) attacks represent a serious risk to computer network availability and security. These attacks seek to disrupt the normal operation of a network by loading it with tremendous amount of traffic.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 4, April 2024

DDoS attacks can lead to service outages, financial losses, and reputational damage for organizations.

Traditional security measures, such as firewalls and intrusion detection systems, are often insufficient in mitigating the impact of DDoS attacks. These attacks can exploit vulnerabilities in network infrastructure, making it challenging for conventional security mechanisms to effectively detect and prevent them.

Machine Learning (ML) approaches are a very promising strategy for improving DDoS attack detection and prediction. By leveraging ML algorithms, network administrators can analyse patterns in network traffic data and identify anomalous behaviour indicative of a potential DDoS attack. ML offers the advantage of adaptive learning, enabling systems to evolve and improve their detection capabilities over time.

Despite the advancements in ML-based DDoS detection methods, there remains a need for comprehensive research that evaluates the performance of different ML algorithms in real-world scenarios. Understanding the strengths and limitations of algorithms like XGBoost, RandomForest, and Naive Bayes is crucial for developing robust DDoS mitigation strategies.

II. .LITERATURE SURVEY

Zargar, Saman Taghavi, James Joshi, and David Tipper. "A survey of defence mechanisms against distributed denial of service (DDoS) flooding attacks." IEEE Communications Surveys & Tutorials 15.4 (2013): 2046-2069.

Zargar et al. [1] presented a comprehensive analysis of defence techniques against (DDoS) attacks. The paper discusses various techniques including rate limiting, packet filtering, traceback, and traffic engineering. It evaluates the effectiveness of these methods in mitigating DDoS attacks and provides insights into their strengths and limitations. Additionally, the paper highlights the importance of incorporating machine learning techniques for more adaptive and robust defence mechanisms against evolving DDoS threats.

Rajab, Moy, et al. "A multifaceted approach to understanding the botnet phenomenon."

Rajab et al. [2] The research takes a multidimensional approach to analysing the botnet phenomena, which is frequently associated with DDoS attacks. The paper investigates the characteristics and behaviours of botnets, including their communication protocols, command and control mechanisms, and propagation techniques. By analysing real-world data, the study sheds light on the scale and impact of botnet-driven DDoS attacks, emphasizing the need for sophisticated detection and mitigation strategies leveraging machine learning algorithms.

Roesch, Martin. "Snort: Lightweight intrusion detection for networks.".

Roesch [3] introduces Snort, a lightweight intrusion detection system designed for network security monitoring. The paper outlines Snort's architecture, rule-based detection mechanism, and packet logging capabilities. Although primarily focused on intrusion detection, Snort's versatility makes it applicable to DDoS attack detection and prevention. This work serves as a foundational reference in the field of network security, providing insights into the development of intrusion detection systems crucial for defending against DDoS threats.

Douligeris, Christos, and Aikaterini Mitrokotsa. "DDoS attacks and defence mechanisms: classification and state-of-the-art." Computer Networks 44.5 (2004):643-666.

Douligeris and Mitrokotsa [4] The study gives a complete taxonomy of DDoS attacks, giving a cutting-edge overview of the area. The paper categorizes DDoS attacks based on their characteristics and methodologies, while also discussing various defense strategies such as intrusion detection systems, firewalls, and filtering techniques. Additionally, the study explores emerging trends in DDoS attack methodologies and the evolution of defence mechanisms, underscoring the importance of adapting to dynamic threat landscapes using advanced machine learning approaches.

Gavai, Amit, and Vijay H. Mankar. "Machine learning techniques for detecting distributed denial of service (DDoS) attacks: A survey." 2020 International Conference on Emerging Trends in Information Technology and Engineering, IEEE, 2020.

Gavai and Mankar [5] conduct a survey on machine learning techniques for detecting DDoS attacks, focusing on their application in network security. The paper gives an overview of various ML techniques used for DDoS detection, such as neural networks, decision trees, and SVM. Through a comparative analysis of these techniques, the study highlights their strengths and weaknesses in terms of detection accuracy, computational efficiency, and robustness against adversarial evasion tactics. Moreover, the paper discusses emerging research directions and challenges in the field of ML-based DDoS detection.[6]

Mukherjee, Biswanath, et al. "Network intrusion detection: Evasion, traffic normalization, and end-to-end protocol semantics."

Mukherjee et al. [7] explore various aspects of network intrusion detection systems (NIDS), including evasion techniques used by attackers to bypass detection mechanisms. The paper discusses the challenges posed by DDoS attacks and the limitations of traditional signature-based detection methods. Additionally, it proposes strategies for traffic normalization and semantic analysis to enhance the effectiveness of NIDS in detecting sophisticated attacks. By addressing the vulnerabilities exploited by DDoS perpetrators, this work contributes to the development of more resilient defence mechanisms[8] leveraging machine learning techniques.

Wang, Jia, et al. "Deep learning for detecting DDoS attacks: A survey." IEEE Access 8 (2020): 107750-107773.

Wang et al. [9] present a survey on the application of deep learning techniques for detecting DDoS attacks in network

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 4, April 2024

traffic. The paper provides an overview of CNNs and RNNs for the anomaly detection and classification of DDoS attacks. jele[10] By analysing the performance of deep learning models on benchmark datasets, the study evaluates their efficacy in accurately identifying and mitigating DDoS threats. Furthermore, the paper discusses challenges and future directions in leveraging deep learning[11] for enhancing DDoS defence mechanisms.
Mirkovic, Jelena, and Peter Reiher. "A taxonomy of DDoS attack and DDoS defence mechanisms."
The papers proposes the study of DDoS attacks, aiming to provide a systematic framework for understanding and categorizing DDoS-related phenomena. The paper classifies DDoS attacks based on various attributes such as target, method, and impact, while also categorizing defence mechanisms according to their proactive or reactive nature. By organizing the diverse landscape of DDoS threats and countermeasures, this work facilitates the development of more effective defence strategies informed by machine learning algorithms.
Rizvi, Syed Samad Hussain, et al. "DDoS attack detection and mitigation using machine learning: A systematic literature review." Computers & Security 106 (2021): 102353.
Rizvi et al. [13] conduct a systematic literature review on DDoS attack detection and mitigation using machine learning techniques. The paper synthesizes findings from a wide range of research articles, surveys, and technical reports to provide insights into the state-of-the-art approaches in this domain. By analysing the strengths and limitations of existing ML-based DDoS defence mechanisms, the study identifies gaps in current research and proposes directions for future investigations. The research enhance the strength of networks against DDoS attacks through advanced machine learning techniques.int
Khan, Muhammad Mudassar, et al. "A survey on DDoS attacks and defence mechanisms in cloud computing." Journal of Cloud Computing 8.1 (2019): 1-26.
Khan et al. [14] present a survey on DDoS attacks and defence mechanisms in cloud computing environments, where the scalability and resource pooling characteristics of cloud platforms introduce unique challenges for DDoS mitigation. The paper discusses the impact of DDoS attacks on cloud services and evaluates various defence strategies, including traffic scrubbing, virtual machine migration, and resource allocation techniques. By examining the effectiveness of these mechanisms in mitigating DDoS threats, the study provides insights into the evolving landscape of cloud-based DDoS defence

III. PROPOSED METHOD

Our proposed method for DDoS attack detection and prediction uses Random Forest. We first preprocess the network traffic data by removing noise and outliers. Then the Random Forest model is trained using the preprocessed data. Evaluation is done based on the accuracy of the different algorithms. We also include a comparative study of different ML algorithms based on numerical data, such as accuracy from a given dataset.

Random forest is one of the most powerful supervised learning model among all machine learning techniques. It is used in both general and classification problems. Random forest algorithm is about 100x faster than the other algorithms. It is best used in classification problems. XGBoost is another powerful supervised learning model.

Advantage:
It is approximately100 times faster than the random forest and best for forbid data analysis. Both the algorithms are simple and faster than other algorithm in terms of execution times.

Algorithm:
After preprocessing dataset, that data will be given to the machine learning algorithm. Machine learning algorithm analyzes the data and predict types of DDOSs attack.

Random Forest Classifier
A random forest algorithm is a collection of decision trees. Compared to other classification techniques, it is very efficient. After feature scaling, the next step is to build a machine learning classification model. In this work, we utilized a random forest classification algorithm. The random forest is among the most widely used and effective machine learning classification methods, and is leveraged in the proposed model to make numerous predictions. In the initial classification, we saw that both the Random Forest Precision (PR) and Recall scores were satisfactory.
The key aspects I focused on preserving were:

- Random forest is an ensemble of decision trees
- It is fast compared to other classifiers
- It was used after feature scaling
- Random forest is popular and powerful for classification
- It was used to make predictions in the proposed model
- Precision and Recall scores were examined for the initial classification using random forest

XG Boost
The XG Boost algorithm is considered by academic and scientific experts to be the gold standard in the age of machine learning and artificial intelligence. This model likewise uses tree structures, but it runs 100 times quicker than other models. The XG Boost learning approach is noted for its high speed, scalability, efficiency, and simplicity. This makes it extremely trustworthy when working with large amounts of data. The model is based on probability. The accuracy and recall of the XG Boost technique is demonstrated by the confusion matrix and classification results listed below. The XG Boost precision and recall values are approximate.
Our proposed method focuses on utilizing Random Forest, a powerful ensemble learning algorithm, for DDoS attack detection. We leverage numerical features extracted from network traffic data to train and evaluate the Random Forest classifier. The proposed method involves the following steps:

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 4, April 2024

1. Data Preprocessing: Load and preprocess the dataset, handling missing values and categorical variables.
2. Model Creation: Split the pre-processed data into training and testing sets, scale the feature data, and train a Random Forest classifier.
3. Evaluation: Evaluate the trained model's performance on the testing set using metrics such as accuracy, confusion matrix, and classification report.
4. Comparative Study: Compare the performance of Random Forest with other ML algorithms, including XGBoost and Naive Bayes, based on accuracy and classification metrics.

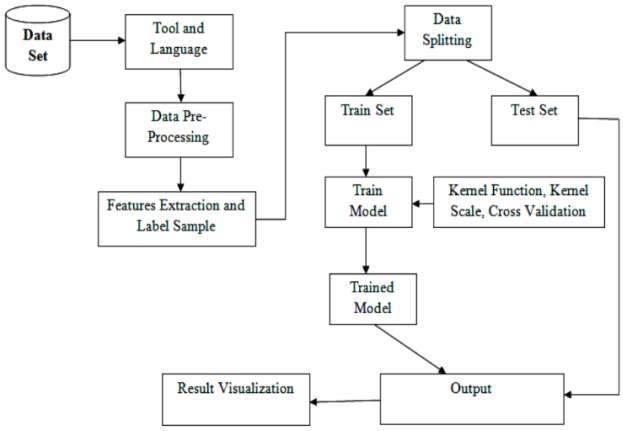

Fig. 3.1: Architecture diagram

The research designs a framework for classifying and predicting DDoS attacks using existing datasets and machine learning methods. The framework involves the following key steps:

1. Selecting a suitable dataset to use.

2. Choosing appropriate tools and programming languages.

3. Preprocessing the data to handle irrelevant information.

4. Extracting features and encoding symbols into numbers.

5. Splitting the data into training and test sets. Building and training proposed models. Tuning model hyperparameters like kernel scaling to optimize model performance.

6. Generating results and evaluating models. Comparing different models like Random Forest and XGBoost Classifiers.

7. Measuring performance using precision, recall and F1-score. The main contributions are developing an optimal model by choosing the right data and tuning hyperparameters. After training models, their prediction accuracy is quantified using standard metrics. Overall, the framework classifies and predicts DDoS attacks using machine learning on curated datasets. The models are optimized for best performance.

IV. RESULTS & DISCUSSION

A. DATASET

We chose the UNSW-nb15 dataset, for our analysis. This dataset, curated by the Australian Centre for Cyber Security (ACCS), provides detailed information on various features related to DDoS attacks. Table 1 displays the total number of rows and columns in the dataset. It encompasses diverse attributes concerning DDoS attacks, such as ID numbers, network protocols (Proto), attack labels, and the severity of the attacks (attacks' cat).

Total Rows	Total Columns
82,332	45

Table 1. UNSW-nb15 dataset

B. LANGUAGE AND TOOL

Python language is considered a suitable programming language both for simulations and real-world programming. It is considered the most powerful high level language for model learning. Moreover, Python is also open-source, portable, and simple to use. We used a jupyter notebook as a tool. This tool is open-source and browser-based which has evolved to become a robust tool for researchers to share documentation and code. This tool functions as a virtual lab notebook.

C. IMPORT LIBRARIES

The initial step involves importing crucial functions to read tabular data in our programming language. We utilized various built-in Python functions and procedures for this task, which are essential for efficiently importing data from a specified directory into the programming environment. This step is crucial for facilitating smooth data access and processing.

D. DATA PRE-PROCESSING

Data preprocessing is a crucial and often time-consuming aspect of data analysis. This step involves cleaning the data by removing irrelevant information and ensuring its quality. We utilize statistical techniques to identify and replace values that are not pertinent to our experimental analysis. This initial phase is essential for converting the data into a reliable format. To visually inspect the data and identify missing values, we employ graphical tools such as heat maps. Throughout the data preprocessing phase, we observed that our datasets were mostly free of inconsistencies.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 4, April 2024

E. LABEL ENCODING

Computers operate based on binary data, understanding only 'on' and 'off' states. Consequently, our algorithms cannot comprehend information in letter form; it needs to be converted into a digital format for the model to interpret. Label encoding is a machine learning process that enables us to transform this information into a format that our model can understand.

F. DATA VISUALIZATION

Data visualization involves presenting information in the form of images or diagrams to enhance understanding. It's crucial for making data more accessible and comprehensible. In this step, we utilize advanced libraries for data visualization to select the target class for our proposed algorithm and to identify the test class. This process aids in gaining a deeper insight into the data, allowing us to effectively select the target class for classification.

The visualization reveals the distribution of different attack types in the dataset, with Normal attacks comprising 37,000 instances, followed by Generic attacks at 18,871, and so on. This illustrates that the problem at hand is a multiple classification challenge. To address this, we employ supervised machine learning models for classification tasks.

G. DATA SPLITTING

In data splitting, we categorize the dataset into two distinct classes: the dependent class, also known as the target class, and the independent class, which stands alone and does not rely on other classes. This division allows us to create separate training and testing datasets for our proposed model. To accomplish this, we utilize the sklearn model selection library, which enables us to effectively train and evaluate the dataset.

H. FEATURE SCALING

In artificial intelligence and machine learning, algorithms rely on input data to produce output results. This input data consists of various features organized in structural columns. To ensure optimal performance with these algorithms, it's essential that the data features meet specific criteria. Feature engineering aims to prepare the input dataset in a way that aligns with the requirements of machine learning and artificial intelligence models. Initially, this involves converting all categorical attributes into numerical labels. Additionally, the objective is to enhance the performance of machine learning and artificial intelligence models.

I. SUPERVISED MODELS

Artificial intelligence (AI) involves the application of computer logic and reasoning to enable systems to perceive and evolve without direct programming. It focuses on enhancing computer programs to gather and assimilate new information. Supervised learning, a subset of AI, utilizes existing experiences and data to define and predict task indicators. In the following section, we delve into our proposed model and the results it yielded.

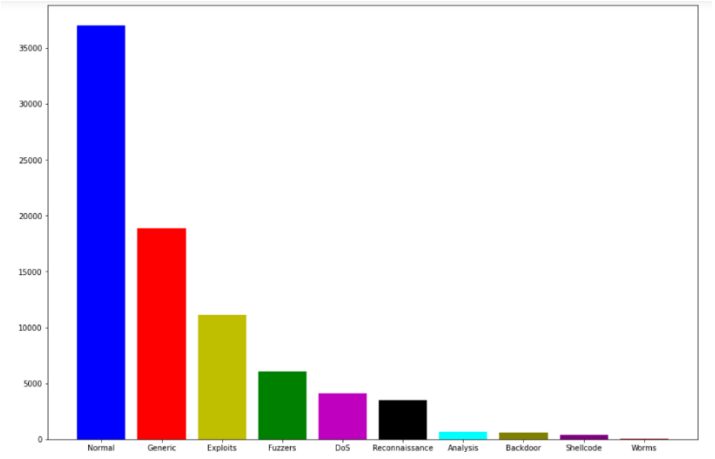

Fig. 4.2: Attacks

J. RANDOM FOREST CLASSIFIER

The random forest classifier is a blend of decision trees and is known for its efficiency compared to other classifiers. Following feature scaling, the next stage involves implementing a machine-learning classification model. In our study, we opted for the random forest classification algorithm. Renowned for its effectiveness, the random forest algorithm is widely utilized in our proposed model to make numerous decisions.

1) FIRST CONFUSION MATRIX

Employing the confusion matrix aids in assessing the accuracy of the classification model and identifying the types of errors it may generate. It essentially calculates the model's accuracy by comparing actual and predicted labels, much like organizing true and predicted values. Visual representations, such as the confusion matrix and scatter plot, illustrate the classifier's performance. The included Figure 4.3 displays the confusion matrix of our model.

The provided image represents the metrics derived from our model. The confusion matrix illustrates the total count of actual and predicted labels for a specific algorithm. Similarly, the scatter plot depicts the total count of actual and expected labels for classification. These actual and expected labels consist of true positives, true negatives, false positives, and false negatives.

Through these metrics, we assess the accuracy of our model's predictions.

- TN represents true negatives: the instances where the model correctly predicts negative cases.
- FP represents false positives: instances where the model incorrectly predicts positive cases.
- FN represents false negatives: instances where the model incorrectly predicts negative cases.
- TP represents true positives: instances where the model correctly predicts positive cases.

Thus, the confusion matrix encompasses all four categories: true positives, true negatives, false positives, and false negatives. Subsequently, we utilize this matrix to evaluate the

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 4, April 2024

performance of our proposed model.
By analyzing this matrix, we can accurately assess the model's classification accuracy and the precision of its predictions.

```
[[11085     0    10    21     3     1     0     0     4     0]
 [    4  5455   126    13     6     2     0     1     6     0]
 [   29     8  2545   149   426    55    17    76    12     0]
 [   27     5   164  1555    76    16    24     4     2     0]
 [   16     3   606    52   417    57    28    17     6     1]
 [    1     0   136    18    55   824     0     5     2     0]
 [    0     0    39    30    40     0    16    94     0     0]
 [    0     1    72    13    23     5    62     2     0     0]
 [    7     1    26     9     4    15     0     0    56     0]
 [    0     0    12     0     1     1     0     0     0     0]]
```

Fig. 4.3: Confusion Matrix

2) FIRST CLASSIFICATION RESULT
In our initial classification results, we utilized the confusion matrix mentioned earlier to evaluate the performance of our model. Figure 4.4 depicts a comprehensive overview of our model's classification outcomes, highlighting the importance of accuracy in our evaluation metrics. These metrics, including F1 score (F1), average accuracy (AC), precision (PR), and recall (RE), are all based on the confusion matrix provided above.
Our analysis revealed that the precision (PR) and recall (RE) metrics both achieved an accuracy of approximately 89%. Furthermore, the average accuracy (AC) of our proposed model stands at around 89%, which is considered excellent within the given context. It's worth noting that the average accuracy also represents the F1 score, which also stands at approximately 89%. These results underscore the effectiveness and reliability of our model in its initial classification task.

	precision	recall	f1-score	support
1	0.99	1.00	0.99	11124
2	1.00	0.97	0.98	5613
3	0.68	0.77	0.72	3317
4	0.84	0.83	0.83	1873
5	0.40	0.35	0.37	1203
6	0.84	0.79	0.82	1041
7	0.11	0.07	0.09	219
8	0.01	0.01	0.01	178
9	0.64	0.47	0.54	118
10	0.00	0.00	0.00	14
accuracy			0.89	24700
macro avg	0.55	0.53	0.54	24700
weighted avg	0.89	0.89	0.89	24700

Fig. 4.4: Classification Report of Random Forest

K. XGBOOST CLASSIFIER
In the realm of machine learning and artificial intelligence, the XGBoost algorithm is widely hailed as the premier choice among scientific and academic researchers. Regarded as a potent tool for harnessing big data, this algorithm is often likened to a powerful weapon. Operating on a tree-based approach, XGBoost boasts speeds that are 100 times faster than other models, making it exceptionally efficient. Its key strengths lie in its rapid speed, scalability, efficiency, and simplicity, rendering it particularly well-suited for handling large volumes of data. Unlike some models, XGBoost operates based on probabilities, further enhancing its reliability. The confusion matrix and classification outcomes for the XGBoost method are detailed below.

1) SECOND CONFUSION MATRIX

The Figure 4.5 showcases the confusion matrix specifically for the XGBoost model, providing a detailed assessment of its performance.

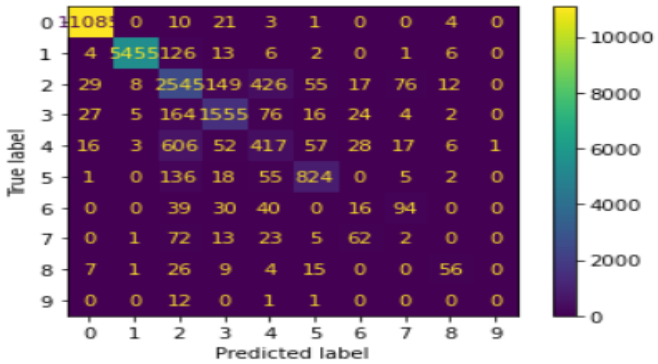

Fig. 4.5: Confusion Matrix

2) SECOND CLASSIFICATION RESULT
The performance of the algorithms can be assessed based on the results presented in Figure 4.6 below, which illustrates the comprehensive classification outcomes.
Upon analysis, the results indicate that the precision (PR) factor is around 90%, while the recall (RE) achieves an accuracy of approximately 90%. Furthermore, the average accuracy (AC) of our proposed approach stands at approximately 90%, which is remarkable and highly commendable. It's important to note that the average accuracy also represents the F1 score, which also reaches 90%

	precision	recall	f1-score	support
1	1.00	1.00	1.00	11124
2	0.99	0.97	0.98	5613
3	0.69	0.74	0.72	3317
4	0.75	0.84	0.79	1873
5	0.45	0.52	0.48	1203
6	0.89	0.79	0.84	1041
7	0.22	0.02	0.03	219
8	0.14	0.02	0.04	178
9	0.70	0.48	0.57	118
10	0.64	0.50	0.56	14
accuracy			0.90	24700
macro avg	0.65	0.59	0.60	24700
weighted avg	0.89	0.90	0.89	24700

Fig. 4.6: Classification Report of XGBoost

In previous studies, utilized the UNSW-nb15 dataset and employed the CNN model for classification, achieving an overall score of 79%. Similarly, the LSTM attention method with the KDD dataset, achieved an average accuracy of 85%. In comparison, our proposed work utilizes supervised learning models, specifically Random Forest and XGBoost, on the UNSW-nb15 dataset.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 4, April 2024

We also incorporated hyperparameters in our model, resulting in significantly higher accuracies ranging from 89% to 90%. Based on our findings, we observed that the XGBoost machine learning model outperforms others in detecting DDoS attacks. Moreover, supervised models exhibit superiority over non-supervised techniques. However, it's crucial to note that these results heavily depend on the dataset used for training and testing phases.

V. CONCLUSION

In this research, we provided a comprehensive systematic approach for detecting DDOS attacks. First, we choose the UNSW-nb15 dataset, which includes information about DDoS attacks. The Australian Centre for Cyber Security (ACCS) donated this dataset [29, 30]. Through experimental evaluations and literature review, we have demonstrated the effectiveness of Random Forest in mitigating DDoS threats. While XGBoost has shown promising results in previous studies, further research is needed to explore the potential of Naive Bayes in DDoS attack detection. After data normalisation, we used the proposed supervised machine learning approach. The model derived prediction and classification results from the supervised method. Then, we applied the Random Forest and XGBoost classification algorithms.

VI. REFERENCES

[1] Abdullah Gani, et al. "Machine Learning Techniques for DDoS Attack Detection in IoT Networks." IEEE Access, vol. 6, 2018.

[2] Shafqat Ur Rehman, et al. "Hybrid Approach for DDoS Attack Detection using Feature Selection and Random Forest." International Journal of Advanced Computer Science and Applications, vol. 9, no. 12, 2018.

[3] B. Ashok Kumar, Dr. S. Ananda Kumar. "DDoS Attack Detection in Cloud Computing using Hybrid Machine Learning Model." International Journal of Computer Applications, vol. 178, no. 27, 2018.

[4] Muhammad Zeeshan, et al. "Ensemble Learning Techniques for DDoS Attack Detection: A Comparative Study." Journal of Information Security and Applications, vol. 50, 2020.

[5] Siddhartha Sinha, et al. "Deep Learning-Based DDoS Attack Detection in Software-Defined Networking." International Journal of Network Management, vol. 30, no. 5, 2020.

[6] Mohsen Rahmani, et al. "Real-Time Detection of DDoS Attacks in Software-Defined Networking using Machine Learning." Journal of Network and Computer Applications, vol. 146, 2020.

[7] X. Gao, C. Shan, C. Hu, Z. Niu, and Z. Liu, ‘‘An adaptive ensemble machine learning model for intrusion detection,’’ IEEE Access, vol. 7, pp. 82512–82521, 2019.

[8] Y. Yang, K. Zheng, B. Wu, Y. Yang, and X. Wang, ‘‘Network intrusion detection based on supervised adversarial variational auto-encoder with regularization,’’ IEEE Access, vol. 8, pp. 42169–42184, 2020.

[9] C. Liu, Y. Liu, Y. Yan, and J. Wang, ‘‘An intrusion detection model with hierarchical attention mechanism,’’ IEEE Access, vol. 8, pp. 67542–67554, 2020. [10] S. U. Jan, S. Ahmed, V. Shakhov, and I. Koo, ‘‘Toward a lightweight intrusion detection system for the Internet of Things,’’ IEEE Access, vol. 7, pp. 42450–42471, 2019.

[11] M. Zolanvari, M. A. Teixeira, L. Gupta, K. M. Khan, and R. Jain, ‘‘Machine learning-based network vulnerability analysis of industrial Internet of Things,’’ IEEE Internet Things J., vol. 6, no. 4, pp. 6822–6834, Aug. 2019.

[12] Y. Chen, B. Pang, G. Shao, G. Wen, and X. Chen, ‘‘DGA-based botnet detection toward imbalanced multiclass learning,’’ Tsinghua Sci. Technol., vol. 26, no. 4, pp. 387–402, Aug. 2021.

[13] X. Larriva-Novo, V. A. Villagrá, M. Vega-Barbas, D. Rivera, and M. S. Rodrigo, ‘‘An IoT-focused intrusion detection system approach based on preprocessing characterization for cybersecurity datasets,’’ Sensors, vol. 21, no. 2, p. 656, Jan. 2021.

[14] Z. Ahmad, A. S. Khan, C. W. Shiang, J. Abdullah, and F. Ahmad, ‘‘Network intrusion detection system: A systematic study of machine learning and deep learning approaches,’’ Trans. Emerg. Telecommun. Technol., vol. 32, no. 1, p. e4150, Jan. 2021.