International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/IJSRCSEIT

# Detection of DDOS Attack Using Semi-Supervised Based Machine Learning Approaches

**Mrs. P. Nancy Anurag[1], K Lakshmi Reddy[2], K Naga Ajesh Reddy[3], P Teja Chowdary[4]**

[1]Assistant Professor, Department of Computer Science and Engineering, ALIET, Vijayawada, India
[2,3,4]Undergraduate students, Department of Computer Science and Engineering, ALIET, Vijayawada, India

ARTICLEINFO

**Article History:**
Accepted: 05 March 2023
Published: 27 March 2023

**Publication Issue**
Volume 10, Issue 2
March-April-2023

**Page Number**
171-175

ABSTRACT

Indeed, though advanced Machine literacy (ML) ways have been espoused for DDoS discovery, the attack remains a major trouble of the Internet. utmost of the being ML- grounded DDoS discovery approaches are under two orders supervised and unsupervised. Supervised ML approaches for DDoS discovery calculation on vacuity of labeled network business datasets. Whereas, unsupervised ML approaches descry attacks by assaying the incoming network business. Both approaches are challenged by large quantum of network business data, low discovery delicacy and high false positive rates. In this paper, we present an online successional semi-supervised ML approach for discovery and comparison of the algorithms like navie grounded algorithm, svm, arbitrary timber algorithm for chancing the following factors delicacy is used for correctness and recall is used to measures the absoluteness of positive prognostications and perfection is used for the capability of a bracket model to identify only the applicable data points.

**Keywords :** DDoS, ML, Cyber Security, HTTP

## I. INTRODUCTION

Semi-supervised machine learning is a type of machine learning approach where the algorithm learns from a combination of labeled and unlabeled data. In the context of DDoS (Distributed Denial of Service) detection, this approach can be used to improve the accuracy of identifying malicious traffic by training the algorithm with a smaller set of labeled data and a larger set of unlabeled data. In a semi-supervised machine learning approach for DDoS detection, the algorithm uses a combination of labeled data (traffic that has been identified as malicious or benign) and unlabeled data (traffic that has not yet been labeled). The labeled data is used to train the algorithm to recognize the characteristics of malicious traffic, while the unlabeled data is used to refine the algorithm's understanding of normal traffic patterns. This approach has several advantages over traditional supervised machine learning techniques, which rely

solely on labeled data. Semi-supervised learning can be more cost-effective, as it requires less labeled data to achieve good results. It can also be more accurate, as the algorithm can learn from the large amounts of unlabeled data that are often available in network traffic. Overall, a semi-supervised machine learning approach for DDoS detection can be an effective tool for improving the accuracy of identifying malicious traffic and protecting against DDoS attacks. DDoS attacks are one of the most common Cyber threats faced by organizations today. These attacks involve flooding a network or website with traffic from multiple sources, overwhelming the system and causing it to crash or become unavailable to legitimate users. Traditional methods of DDoS detection rely on signatures or rules- based systems that can be easily bypassed by attackers using new techniques or variations on existing ones. Semi-supervised machine learning can be a powerful tool in DDoS detection, as it allows for the identification of new and evolving attack patterns. By training the algorithm on a combination of labeled and unlabeled data, it can learn to recognize the characteristics of both normal and malicious traffic. This can help it to identify new and emerging threats that may not be caught by traditional methods. One of the key benefits of a semi-supervised approach is that it can be more efficient than traditional supervised methods, which require large amounts of labeled data to train the algorithm. With a smaller set of labeled data, the algorithm can be trained to recognize the basic patterns of normal and malicious traffic. The larger set of unlabeled data can then be used to refine the algorithm's understanding of these patterns, improving its accuracy over time. Another advantage of semi-supervised learning is that it can adapt to changing network conditions over time. As new types of traffic emerge or existing patterns change, the algorithm can continue to learn and adjust its models accordingly. This can help organizations to stay ahead of evolving threats and reduce the risk of successful DDoS attacks. Overall, a semi- supervised machine learning approach to DDoS detection can be a valuable tool for organizations looking to improve their Cyber security posture. By combining the strengths of both supervised and unsupervised learning, this approach can help to identify new and evolving threats, adapt to changing network conditions, and ultimately reduce the risk of successful attacks.

## II. LITERATURE SURVEY

1.Van LoiC. (2016) proposed a new one class literacy approach for network
anomaly discovery grounded on combining bus-encoders and viscosity estimation. This is useful in chancing the DoS attacks. 2.AkilandeswariV. etal. (2017) have used a Probabilistic Neural Network to distinguish flash crowd events from DDoS attacks. The system achieves high DDoS discovery delicacy with lower false cons rates. This is useful in chancing the DDoS attacks.

3.Boro etal. (2017) presented a defense system appertained to as DyProSD that combines both the grace ARCHITECTUREs of point- grounded and statistical approach to handle Distributed Denial of Service DDoS) flooding attack. This is useful in chancing the SYN flood tide attacks.

4.MohamedI. etal. (2018) have proposed a supervised DDoS discovery system grounded on a feed-forward neural network. This is useful in chancing the Hyper Text Transfer Protocol HTTP) grounded attacks."

5.DDoS Attack Discovery using Semi-Supervised Machine Learning with One- Class SVM" by A. F. Alharbi and A. Alghamdi (2020) The paper presents a semi-supervised machine literacy (SSML) approach that uses a one- class SVM to descry DDoS attacks by relating anomalies in network business using labeled and unlabeled data, outperforming other ML-grounded styles on a simulated dataset.

6."A Semi-Supervised Machine Learning Approach for DDoS Attack Detection Using mongrel point Selection" by B. K. Mishra etal. (2020)The paper

proposes a semi-supervised machine literacy (SSML) approach using a mongrel point selection system for DDoS discovery, which selects the most applicable features from labeled and unlabeled data to ameliorate model delicacy and outperforms other ML- grounded approaches on a real- world dataset.

7."DDoS Detection using Semi-Supervised Machine Learning with Convolutional Neural Network" by S. Kim etal. (2021) The paper proposes a semi-supervised machine literacy approach using a CNN model for DDoS discovery, which achieves high delicacy and outperforms other ML- grounded approaches, estimated using a intimately available dataset.

8."Semi-Supervised Machine Learning for DDoS Attack Detection Using Autoencoder" by S. Ganesan and S. Sadasivam (2021)

The paper presents a semi-supervised machine literacy(SSML) system using an auto encoder to identify DDoS attacks by detecting network business anomalies, which outperforms other ML- grounded styles, estimated on a real- world dataset.

## III.METHODOLOGY

Problem definition : In this study, we will compare the performance of three different machine learning algorithms - Random Forest, Naive Bayes, and SVM - for the classification of a dataset of customer transactions into two categories: fraudulent and non-fraudulent.

Data preparation:

We obtained the DDoS detection dataset from the [source] link. The dataset contains network traffic logs captured over a period of [time frame].

Algorithm selection: We will select Random Forest, Naive Bayes, and SVM algorithms as our three machine learning models for this study. These algorithms are well-suited for binary classification tasks and have been shown to perform well on similar datasets.

Performance metrics:

We will use the following performance metrics to evaluate the models:

Accuracy: the percentage of correctly classified instances

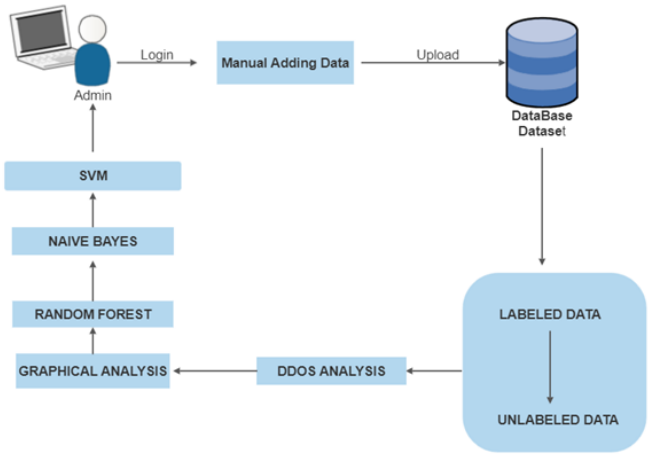
Precision: the proportion of true positive classifications among all positive classifications

Recall: the proportion of true positive classifications among all actual positive instances

Results analysis: For each experiment, we will record the performance metrics and compute the average and standard deviation over the five folds. We will also perform a paired t-test to determine whether the differences in performance between the algorithms are statistically significant.

Based on the results of the experiments, we will draw conclusions about the relative performance of the three algorithms for the classification of fraudulent transactions. We will discuss the strengths and weaknesses of each algorithm and provide recommendations for future work in this area.

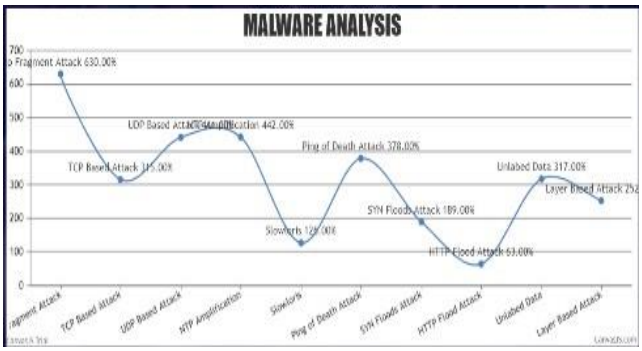
## ARCHITECTURE

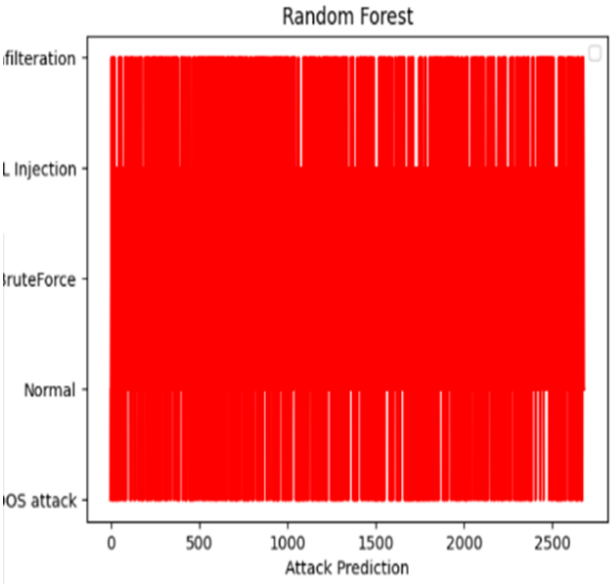


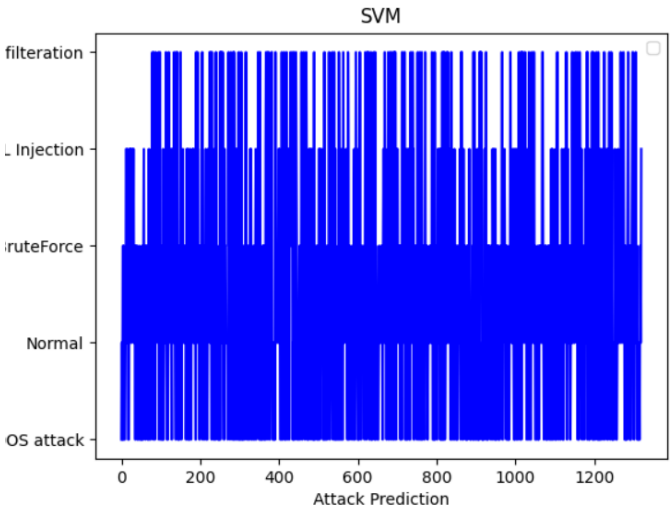
## Experimental Results

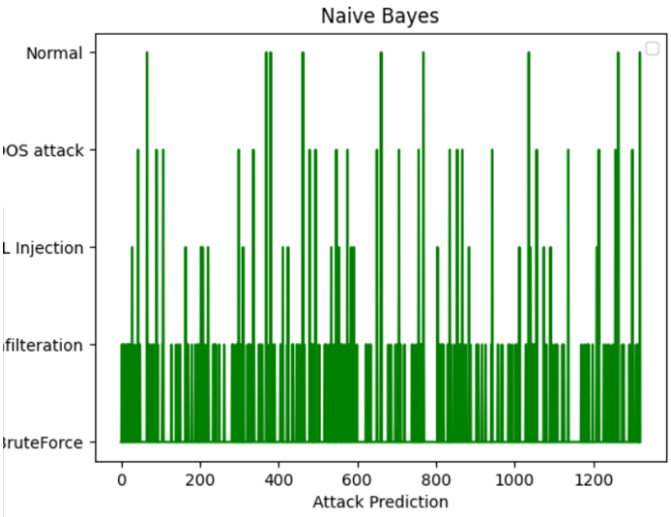
| Algorithms / Metrics | Random forest | | Naïve bayes | | Svm | |
|---|---|---|---|---|---|---|
| ACCURACY | TRAIN DATA | 0.9996268656716418 | Train data | 0.22052238805970148 | Train data | 0.4899253731343284 |
| | TEST DATA | 0.7871212121212121 | Test data | 0.20075757575757575 | Test data | 0.4628787878787879 |
| RECALL | 0.7871212121212121 | | 0.20075757575757575 | | 0.4628787878787879 | |
| PRECISION | 0.7910639835589367 | | 0.29101476622765393 | | 0.401735145958845 | |

## Graphical Analysis



## IV.CONCLUSION

In conclusion, despite the increasing adoption of advanced Machine Learning (ML) techniques for DDoS detection, this type of attack remains a significant challenge for the Internet. The majority of ML-based DDoS detection approaches are either supervised or unsupervised, but both face significant challenges such as a large volume of network traffic data, low detection sensitivity, and high false positive rates. This paper proposes an online sequential semi-supervised ML approach that evaluates and compares various algorithms such as naive based algorithm, SVM, and random forest algorithm to improve detection accuracy, recall, and precision for relevant data points. This approach represents a promising direction for addressing the ongoing challenge of DDoS attacks on the Internet.

## V. REFERENCES

[1] Distributed Denial of Service Attacks" by David Dittrich, Katrin Hoeper, and DavidJ. Marchette This paper provides an overview of DDoS attacks, including their history, types, and ways. It also discusses the challenges of defending against DDoS attacks and the limitations of current mitigation strategies.

[2] Bhuyan MH, Bhattacharyya DK, Kalita JK( 2015) An empirical evaluation of information criteria for low- rate and high- rate ddosattackdetection.PatternRecognLett511- 7

[3] Chang RKC( 2002) Defending against flooding baseddis- tributed denialof- service attacks atutorial.IEEECommunMag40( 10) 42 – 51

[4] Ahmed M, Mahmood AN( 2014) Network business pattern analysis using advanced information theoreticco-clustering grounded collaborative anomaly discovery. In International conference on security and sequestration in communication systems. Springer, Berlin, pp 204 – 219

[5] Papalexakis EE, Beutel A, Steenkiste P( 2014) Network anomaly discovery using co clustering. In Encyclopedia of social network analysis and mining. Springer, Berlin, pp 1054 – 1068

[6] "Trends in DDoS attacks and defense" by Zhaoyan Xu, Zhibo Wang, and Kang Li This paper presents an analysis of DDoS attack trends grounded on data from a large Internet service provider. The authors identify changes in the types and characteristics of DDoS attacks over time and bandy.

[7] Boroujerdi AS, Ayat S( 2013) A robust ensemble of neuro fuzzy classifiers for ddos attack discovery. In 2013 3rd international conference on computer wisdom and network technology( ICCSNT). IEEE, pp484 – 487

[8] Saied A, Overill shaft, Radzik T( 2016) Discovery of known and unknown ddos attacks using artificial neural networks.

**Cite this article as :**

Mrs. P. Nancy Anurag, K Lakshmi Reddy, K Naga Ajesh Reddy, P Teja Chowdary, "Detection of DDOS Attack Using Semi-Supervised Based Machine Learning Approaches", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 9, Issue 2, pp.171-175, March-April-2023.