International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307

OPEN ACCESS

Available Online at : www.ijsrcseit.com

doi : https://doi.org/10.32628/CSEIT2410223

Voting Model Strategies for Reliable Categorical IoT-DDoS Attack Prediction

Shivani Sinha[1], Dr. Sheshang Degadwala[2]

[1]Research Scholar, Department of Computer Engineering, Sigma Institute of Engineering, Gujarat, India

[2]Professor & Head of Department, Department of Computer Engineering, Sigma University, Gujarat, India

ARTICLEINFO

Article History:

Accepted: 15 March 2024

Published: 28 March 2024

Publication Issue

Volume 10, Issue 2

March-April-2024

Page Number

300-307

ABSTRACT

This research focuses on developing reliable categorical IoT-DDoS attack prediction models using ensemble voting strategies. The study explores various machine learning algorithms suitable for categorical data analysis, employing feature engineering techniques to preprocess IoT data. Ensemble learning methodologies, including bagging, boosting, and stacking, are then utilized to build robust prediction models. Evaluation metrics such as precision, recall, F1-score, and AUC-ROC are used to assess model performance, demonstrating the effectiveness of ensemble voting models in reliably predicting IoT-DDoS attacks. Comparative analyses with individual classifiers highlight the advantages of ensemble approaches in terms of predictive accuracy and robustness against data imbalances and noise. This work contributes to advancing IoT security by providing a practical framework for deploying predictive models that aid in early detection and mitigation of DDoS attacks, enhancing overall resilience against cyber threats in IoT ecosystems.

Keywords :— IoT-DDoS, Ensemble Voting Strategies, Categorical Data Analysis, Machine Learning Algorithms, Feature Engineering.

I. INTRODUCTION

The rapid expansion of the Internet of Things (IoT) has revolutionized the way devices interact and communicate, ushering in an era of unprecedented connectivity and convenience. However, this proliferation of interconnected devices has also introduced significant challenges, particularly in terms of cybersecurity. One of the most pressing concerns is the susceptibility of IoT networks to Distributed Denial of Service (DDoS) attacks, which can disrupt services, compromise data integrity, and jeopardize system availability.

To effectively combat these evolving threats, accurate and timely prediction of IoT-DDoS attacks is imperative. Traditional machine learning techniques have been employed for anomaly detection and attack prediction in IoT environments. However, these methods often struggle to handle the categorical nature of IoT data, which comprises diverse features and behaviors.

In response to these challenges, this paper proposes a novel approach centered on ensemble voting strategies for reliable categorical IoT-DDoS attack prediction. Ensemble learning techniques, such as bagging, boosting, and stacking, offer a promising avenue for improving prediction accuracy and robustness in complex categorical datasets. By leveraging the collective intelligence of multiple classifiers, ensemble voting models can effectively distinguish between normal IoT traffic and malicious DDoS attack patterns.

This research aims to bridge the gap between traditional machine learning approaches and the unique characteristics of IoT data, providing a comprehensive framework for developing predictive models tailored to IoT-DDoS attack scenarios. The study will explore the suitability of different machine learning algorithms for categorical data analysis, investigate feature engineering techniques to enhance data preprocessing, and evaluate the performance of ensemble voting models using real-world IoT datasets. The outcomes of this research are expected to contribute significantly to the field of IoT security by offering practical insights into proactive defense strategies and early threat detection mechanisms. By leveraging advanced machine learning methodologies, IoT ecosystems can bolster their resilience against cyber threats, safeguarding critical infrastructure and ensuring the uninterrupted operation of IoT-enabled services.

II. LITERATURE STUDY

Bhayo et al. [1] proposed a machine learning-based framework for detecting DDoS attacks in Software-Defined IoT (SD-IoT) networks. Their study underscores the critical need for advanced detection mechanisms in addressing the evolving landscape of cyber threats targeting IoT infrastructures. By leveraging machine learning algorithms, their framework aims to enhance the resilience of SD-IoT networks against malicious activities, contributing to the overall security posture of IoT ecosystems.

De Lima Filho et al. [2] presented an online approach for DoS/DDoS attack detection using machine learning techniques. The study emphasizes the importance of real-time detection capabilities to mitigate the impact of disruptive attacks on IoT devices and networks. By implementing machine learning algorithms for continuous monitoring and analysis, their approach seeks to bolster the security resilience of IoT environments in the face of dynamic threats.

Jan et al. [3] focused on developing a lightweight intrusion detection system specifically tailored for IoT deployments. Their research addresses the inherent challenges posed by resource-constrained IoT devices while emphasizing the criticality of effective intrusion detection mechanisms. By adopting lightweight algorithms and efficient data processing techniques, their system aims to provide proactive security measures without imposing significant overhead on IoT devices.

Otoum and Nayak [4] introduced AS-IDS, an anomaly and signature-based Intrusion Detection System (IDS) designed for IoT environments. Their approach combines anomaly detection for novel threats and signature-based detection for known attack patterns, offering comprehensive security coverage. By integrating adaptive detection mechanisms, AS-IDS demonstrates the versatility required to combat the diverse and evolving nature of IoT-related security threats.

Zagrouba and AlHajri [5] delved into machine learning-based attacks detection and countermeasures in IoT, underscoring the role of machine learning in enhancing security postures. Their research contributes to the development of effective strategies for detecting and mitigating threats targeting IoT infrastructures, thereby improving overall cybersecurity resilience.

Mohanta et al. [6] conducted a survey on IoT security challenges and solutions, emphasizing the intersection of machine learning, artificial intelligence, and blockchain technology in addressing security concerns. Their work provides a comprehensive

overview of the current landscape of IoT security, identifying key challenges and proposing innovative solutions.

Pokhrel et al. [7] explored botnet detection in IoT using machine learning, highlighting the significance of proactive measures against sophisticated threats. Their research sheds light on the potential vulnerabilities posed by botnet attacks in IoT environments and offers insights into effective detection strategies.

Anthi et al. [8] developed a supervised intrusion detection system for smart home IoT devices, showcasing targeted security solutions for specific IoT applications. Their study focuses on enhancing security measures for IoT devices commonly used in smart home environments, addressing unique security challenges within this context.

Essaid et al. [9] proposed a collaborative DDoS mitigation solution based on Ethereum smart contracts and RNN-LSTM, demonstrating innovative approaches to mitigating DDoS attacks in IoT environments. Their research integrates blockchain technology with machine learning techniques to develop a robust defense mechanism against DDoS threats.

Javaid et al. [10] investigated the use of blockchain technology for mitigating DDoS attacks, showcasing the potential of decentralized architectures in enhancing resilience against attacks. Their work contributes to the exploration of novel security paradigms that leverage blockchain's inherent properties for mitigating DDoS risks.

Singh et al. [11] utilized blockchain for mitigating distributed denial of service (DDoS) attacks, highlighting the potential of blockchain in enhancing security measures. Their research emphasizes the importance of leveraging emerging technologies to develop robust defense mechanisms against evolving cyber threats.

Bhardwaj et al. [12] proposed IoT-DDoS prevention using edge computing, highlighting the role of edge resources in augmenting security measures. Their study explores the feasibility of leveraging edge computing capabilities to enhance the detection and mitigation of DDoS attacks targeting IoT infrastructures.

Kumar et al. [13] developed a distributed framework for detecting DDoS attacks in smart contract-based Blockchain-IoT systems, integrating fog computing for enhanced security. Their research demonstrates a holistic approach to DDoS mitigation by combining blockchain technology with fog computing for improved threat detection and response.

Dao et al. [14] focused on securing heterogeneous IoT with intelligent DDoS attack behavior learning, showcasing adaptive approaches to threat mitigation. Their study emphasizes the importance of intelligent learning techniques in detecting and responding to DDoS attacks targeting diverse IoT devices and protocols.

Adat and Gupta [15] presented a DDoS attack mitigation framework for IoT, emphasizing comprehensive security frameworks tailored to IoT infrastructures. Their research contributes to the development of robust defense mechanisms that address the unique challenges posed by DDoS attacks in IoT environments.

III.PROPOSED SYSTEM

The flow diagram outlines the process of developing a predictive model for detecting IoT-DDoS attacks using machine learning techniques. Here's a detailed description of each step in the diagram:

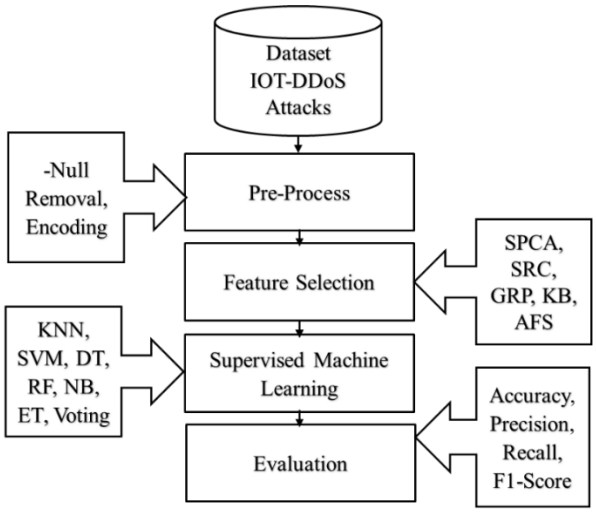

Figure 1. Proposed System

Dataset:

This block represents the initial step in the process, where the dataset containing information related to IoT-DDoS attacks is obtained. This dataset is crucial as it serves as the foundation for training and testing the predictive model. It likely includes features such as network traffic patterns, device behavior, timestamps, and attack labels that indicate whether an instance is a DDoS attack or not.

Pre-Process:

In this block, the dataset undergoes preprocessing steps to ensure data quality and compatibility with machine learning algorithms. Null values or missing data points are identified and either removed or imputed, depending on the extent of missingness. Categorical variables are encoded into numerical format using techniques like one-hot encoding or label encoding.

Feature Selection:

This block involves selecting the most relevant features from the dataset to improve model performance and reduce dimensionality. Techniques such as Sparse Principal Component Analysis (SPCA), Supervised Recursive Feature Elimination (SRC), Genetic Regulatory Programming (GRP), K-Best (KB), and Average-Feature Selection (AFS) are employed to identify features that contribute significantly to predicting IoT-DDoS attacks.

Supervised Machine Learning:

Here, various supervised machine learning algorithms are applied to train the predictive model using the selected features from the previous step. Algorithms such as K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), Naive Bayes (NB), and Extra Trees (ET) are utilized to classify instances and make predictions about whether they represent DDoS attacks.

Voting of [DT, RF, ET]:

This block involves ensemble learning, where predictions from individual decision tree (DT), random forest (RF), and extra trees (ET) models are combined through a voting mechanism. Ensemble voting helps improve prediction accuracy by leveraging the diverse perspectives of multiple models.

Evaluation:

Finally, the model's performance is evaluated using metrics such as accuracy, precision, recall, and F1-score. These metrics provide insights into how well the model is able to correctly classify instances as either normal IoT traffic or DDoS attacks. High values for these metrics indicate a reliable and effective predictive model for IoT-DDoS attack detection.

IV.RESULT ANALYSIS

The dataset introduced is named CICIoT2023. It is a real-time dataset specifically designed and curated for assessing large-scale attacks within IoT environments. The dataset aims to provide a comprehensive benchmark for evaluating the performance of machine learning models and detection algorithms in handling cybersecurity threats targeting IoT devices. The dataset likely contains a diverse range of features related to network traffic, device behavior, attack patterns, and other relevant parameters essential for training and testing predictive models. This initiative contributes significantly to the research community by offering a standardized and realistic dataset that facilitates advancements in IoT security and defense mechanisms. The submission of this dataset to the Journal of Sensors indicates a rigorous validation process and ensures its quality and suitability for research purposes.

Link:
https://www.kaggle.com/datasets/subhajournal/iotintrusion

	flow_duration	Header_Length	Protocol Type	Duration	Rate	Srate	Dra
0	0.000000	54.00	6.00	64.00	0.329807	0.329807	
1	0.000000	57.04	6.33	64.00	4.290556	4.290556	
2	0.000000	0.00	1.00	64.00	33.396799	33.396799	
3	0.328175	76175.00	17.00	64.00	4642.133010	4642.133010	
4	0.117320	101.73	6.11	65.91	6.202211	6.202211	
...	...	...	...	...	...	...	
1048570	1.391925	108.00	6.00	64.00	1.437685	1.437685	
1048571	0.000000	2.14	46.70	65.91	0.000000	0.000000	
1048572	0.132971	30847.00	17.00	64.00	5978.034950	5978.034950	
1048573	0.000000	54.00	6.00	64.00	25.672981	25.672981	
1048574	128.443556	4264.30	7.10	98.80	13.640648	13.640648	

1048575 rows × 47 columns

Figure 2. Dataset Reading

	flow_duration	Header_Length	Protocol Type	Duration	Rate	Srate	Drat
0	0.000000	0.00	1.00	64.00	4.475459	4.475459	0
1	0.000000	0.00	1.00	64.00	3.053283	3.053283	0
2	0.015796	8.21	1.21	64.53	59.344342	59.344342	0
3	0.010202	17.78	1.10	64.00	36.765677	36.765677	0
4	0.000000	0.00	1.00	64.00	3.005904	3.005904	0
...	...	...	...	...	...	...	
1495	158.952476	37496.78	6.83	92.13	52.782361	52.782361	0
1496	45.475748	4584.10	6.48	61.12	0.479950	0.479950	0
1497	7.095029	415266.79	5.96	84.52	129.443848	129.443848	0
1498	20.909274	445858.82	6.22	88.81	38.287177	38.287177	0
1499	1.584323	1301.07	5.88	62.64	2.913833	2.913833	0

1500 rows × 47 columns

Figure 3. Pre-Processing

	SPCA	SRP	GRP	KB	AFS
10	0.785000	0.828333	0.820000	0.663333	0.826667
20	0.829167	0.824167	0.820833	0.814167	0.834167
30	0.834167	0.839167	0.825000	0.832500	0.835000

Figure 4. Fearure Selection Methods

K-Nearest Neighbors

```
[[19  0  0  0  0  0  0  1  0  0  0  0  0  0  0]
 [ 0 20  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0 20  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  1  0 18  0  0  0  1  0  0  0  0  0  0  0]
 [ 0  0  0  0 20  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0 19  0  1  0  0  0  0  0  0  0]
 [ 0  0  1  0  0  0 19  0  0  0  0  0  0  0  0]
 [ 0  1  0  0  0  0  0 16  0  0  0  3  0  0  0]
 [ 0  0  0  0  0  0  0  0 20  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0 20  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0 19  0  0  0  1]
 [ 0  0  0  0  0  0  1  2  0  0  0 17  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 20  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0 20  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0 20]]
```

	precision	recall	f1-score	support
DDoS-ACK_Fragmentation	1.00	0.95	0.97	20
DDoS-HTTP_Flood	0.91	1.00	0.95	20
DDoS-ICMP_Flood	0.95	1.00	0.98	20
DDoS-ICMP_Fragmentation	1.00	0.90	0.95	20
DDoS-PSHACK_Flood	1.00	1.00	1.00	20
DDoS-RSTFINFlood	1.00	0.95	0.97	20
DDoS-SYN_Flood	0.95	0.95	0.95	20
DDoS-SlowLoris	0.76	0.80	0.78	20
DDoS-SynonymousIP_Flood	1.00	1.00	1.00	20
DDoS-TCP_Flood	1.00	1.00	1.00	20
DDoS-UDP_Flood	1.00	0.95	0.97	20
DDoS-UDP_Fragmentation	0.85	0.85	0.85	20
DoS-SYN_Flood	1.00	1.00	1.00	20
DoS-TCP_Flood	1.00	1.00	1.00	20
DoS-UDP_Flood	0.95	1.00	0.98	20
accuracy			0.96	300
macro avg	0.96	0.96	0.96	300
weighted avg	0.96	0.96	0.96	300

Figure 5. K-Nearest Neighbour

SVM

```
[[ 0  1  0 19  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 18  0  2  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 20  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  1  0 19  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 20  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 20  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 20  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 15  0  4  0  0  0  0  0  0  0  1  0  0  0]
 [ 0 20  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 20  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 20  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  5  0 15  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 20  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 20  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 20  0  0  0  0  0  0  0  0  0  0  0  0  0]]
```

	precision	recall	f1-score	support
DDoS-ACK_Fragmentation	0.00	0.00	0.00	20
DDoS-HTTP_Flood	0.07	0.90	0.14	20
DDoS-ICMP_Flood	0.00	0.00	0.00	20
DDoS-ICMP_Fragmentation	0.32	0.95	0.48	20
DDoS-PSHACK_Flood	0.00	0.00	0.00	20
DDoS-RSTFINFlood	0.00	0.00	0.00	20
DDoS-SYN_Flood	0.00	0.00	0.00	20
DDoS-SlowLoris	0.00	0.00	0.00	20
DDoS-SynonymousIP_Flood	0.00	0.00	0.00	20
DDoS-TCP_Flood	0.00	0.00	0.00	20
DDoS-UDP_Flood	0.00	0.00	0.00	20
DDoS-UDP_Fragmentation	0.00	0.00	0.00	20
DoS-SYN_Flood	0.00	0.00	0.00	20
DoS-TCP_Flood	0.00	0.00	0.00	20
DoS-UDP_Flood	0.00	0.00	0.00	20
accuracy			0.12	300
macro avg	0.03	0.12	0.04	300
weighted avg	0.03	0.12	0.04	300

Figure 6. Support Vectore Machine

Decision Tree

```
[[19  0  0  0  1  0  0  0  0  0  0  0  0  0  0]
 [ 0 19  0  0  1  0  0  0  0  0  0  0  0  0  0]
 [ 0  0 20  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  1  0 19  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0 20  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0 20  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0 20  0  0  0  0  0  0  0  0]
 [ 0  4  0  2  0  0  0 14  0  0  0  0  0  0  0]
 [ 0  1  0  0  0  0  0  0 19  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0 20  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0 20  0  0  0  0]
 [ 0  0  0  0  0  0  0  2  0  0  0 18  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 20  0  0]
 [ 0  0  0  0  0  0  0  1  0  0  0  0  0 19  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0 20]]
```

	precision	recall	f1-score	support
DDoS-ACK_Fragmentation	1.00	0.95	0.97	20
DDoS-HTTP_Flood	0.76	0.95	0.84	20
DDoS-ICMP_Flood	1.00	1.00	1.00	20
DDoS-ICMP_Fragmentation	0.90	0.95	0.93	20
DDoS-PSHACK_Flood	0.91	1.00	0.95	20
DDoS-RSTFINFlood	1.00	1.00	1.00	20
DDoS-SYN_Flood	1.00	1.00	1.00	20
DDoS-SlowLoris	0.82	0.70	0.76	20
DDoS-SynonymousIP_Flood	1.00	0.95	0.97	20
DDoS-TCP_Flood	1.00	1.00	1.00	20
DDoS-UDP_Flood	1.00	1.00	1.00	20
DDoS-UDP_Fragmentation	1.00	0.90	0.95	20
DoS-SYN_Flood	1.00	1.00	1.00	20
DoS-TCP_Flood	1.00	0.95	0.97	20
DoS-UDP_Flood	1.00	1.00	1.00	20
accuracy			0.96	300
macro avg	0.96	0.96	0.96	300
weighted avg	0.96	0.96	0.96	300

Figure 7. Decision Tree

Naive Bayes

```
[[19  1  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  9  3  0  0  0  8  0  0  0  0  0  0  0  0]
 [ 0  0 19  0  0  0  0  0  0  0  1  0  0  0  0]
 [ 1  1  0 17  0  0  0  1  0  0  0  0  0  0  0]
 [ 0  0  0  0 20  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0 19  1  0  0  0  0  0  0  0  0  0]
 [ 0  0 18  0  0  0  1  0  0  0  1  0  0  0  0]
 [ 0  4  0  0  0  0  1 11  0  0  4  0  0  0  0]
 [ 0  0  0  0  0  0  0  0 20  0  0  0  0  0  0]
 [ 0  0 16  0  0  0  0  0  0  3  1  0  0  0  0]
 [ 0  0  0  0  0  0  0  1  0  0 19  0  0  0  0]
 [ 0  1  0  0  0  0  0  1  0  0  0 18  0  0  0]
 [ 0  0 20  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  1 17  0  0  0  2  0  0  0  0  0  0  0  0]
 [ 0  1  2  0  0  0  0  1  0  0  3  0  0  0 13]]
```

	precision	recall	f1-score	support
DDoS-ACK_Fragmentation	0.95	0.95	0.95	20
DDoS-HTTP_Flood	0.50	0.45	0.47	20
DDoS-ICMP_Flood	0.20	0.95	0.33	20
DDoS-ICMP_Fragmentation	1.00	0.85	0.92	20
DDoS-PSHACK_Flood	0.51	1.00	0.68	20
DDoS-RSTFINFlood	1.00	0.05	0.10	20
DDoS-SYN_Flood	0.08	0.05	0.06	20
DDoS-SlowLoris	0.73	0.55	0.63	20
DDoS-SynonymousIP_Flood	1.00	1.00	1.00	20
DDoS-TCP_Flood	1.00	0.15	0.26	20
DDoS-UDP_Flood	0.66	0.95	0.78	20
DDoS-UDP_Fragmentation	1.00	0.90	0.95	20
DoS-SYN_Flood	0.00	0.00	0.00	20
DoS-TCP_Flood	0.00	0.00	0.00	20
DoS-UDP_Flood	1.00	0.65	0.79	20
accuracy			0.57	300
macro avg	0.64	0.57	0.53	300
weighted avg	0.64	0.57	0.53	300

Figure 9. Naïve Bayes

Random Forest

```
[[19  1  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 19  0  0  0  1  0  0  0  0  0  0  0  0  0]
 [ 0  0 20  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  1  0 18  0  0  0  1  0  0  0  0  0  0  0]
 [ 0  0  0  0 20  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0 20  0  0  0  0  0  0  0  0  0]
 [ 0  1  0  0  0  0 19  0  0  0  0  0  0  0  0]
 [ 0  2  0  0  0  0  0 18  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  4  0 16  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0 20  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0 20  0  0  0  0]
 [ 0  0  0  0  0  0  0  2  0  0  0 18  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 20  0  0]
 [ 0  0  0  0  0  0  0  1  0  0  0  0  0 19  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0 20]]
```

	precision	recall	f1-score	support
DDoS-ACK_Fragmentation	1.00	0.95	0.97	20
DDoS-HTTP_Flood	0.79	0.95	0.86	20
DDoS-ICMP_Flood	1.00	1.00	1.00	20
DDoS-ICMP_Fragmentation	1.00	0.90	0.95	20
DDoS-PSHACK_Flood	1.00	1.00	1.00	20
DDoS-RSTFINFlood	0.95	1.00	0.98	20
DDoS-SYN_Flood	0.83	0.95	0.88	20
DDoS-SlowLoris	0.82	0.90	0.86	20
DDoS-SynonymousIP_Flood	1.00	0.80	0.89	20
DDoS-TCP_Flood	1.00	1.00	1.00	20
DDoS-UDP_Flood	1.00	1.00	1.00	20
DDoS-UDP_Fragmentation	1.00	0.90	0.95	20
DoS-SYN_Flood	1.00	1.00	1.00	20
DoS-TCP_Flood	1.00	0.95	0.97	20
DoS-UDP_Flood	1.00	1.00	1.00	20
accuracy			0.95	300
macro avg	0.96	0.95	0.95	300
weighted avg	0.96	0.95	0.95	300

Figure 8. Random Forest

ExtraTreesClassifier

```
[[19  1  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 17  0  0  1  0  0  0  0  0  0  0  2  0  0]
 [ 0  0 20  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0 19  0  0  0  1  0  0  0  0  0  0  0]
 [ 0  0  0  0 20  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0 20  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0 19  0  1  0  0  0  0  0  0]
 [ 0  4  0  0  0  0  0 16  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0 20  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0 20  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0 20  0  0  0  0]
 [ 0  0  0  0  0  0  0  1  0  0  0 18  0  0  1]
 [ 0  0  0  0  0  0  4  0  0  0  0  0 16  0  0]
 [ 0  0  0  0  0  0  0  1  0  0  0  0  0 19  0]
 [ 0  0  0  0  0  0  0  0  0  0  9  0  0  0 11]]
```

	precision	recall	f1-score	support
DDoS-ACK_Fragmentation	1.00	0.95	0.97	20
DDoS-HTTP_Flood	0.77	0.85	0.81	20
DDoS-ICMP_Flood	1.00	1.00	1.00	20
DDoS-ICMP_Fragmentation	1.00	0.95	0.97	20
DDoS-PSHACK_Flood	0.95	1.00	0.98	20
DDoS-RSTFINFlood	1.00	1.00	1.00	20
DDoS-SYN_Flood	0.83	0.95	0.88	20
DDoS-SlowLoris	0.84	0.80	0.82	20
DDoS-SynonymousIP_Flood	0.95	1.00	0.98	20
DDoS-TCP_Flood	1.00	1.00	1.00	20
DDoS-UDP_Flood	0.69	1.00	0.82	20
DDoS-UDP_Fragmentation	1.00	0.90	0.95	20
DoS-SYN_Flood	0.89	0.80	0.84	20
DoS-TCP_Flood	1.00	0.95	0.97	20
DoS-UDP_Flood	0.92	0.55	0.69	20
accuracy			0.91	300
macro avg	0.92	0.91	0.91	300
weighted avg	0.92	0.91	0.91	300

Figure 10. Extra Tree

```
VotingClassifier
[[19  1  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 19  0  0  1  0  0  0  0  0  0  0  0  0]
 [ 0  0 20  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  1  0 18  0  0  0  1  0  0  0  0  0  0]
 [ 0  0  0  0 20  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0 20  0  0  0  0  0  0  0  0]
 [ 0  1  0  0  0  0 19  0  0  0  0  0  0  0]
 [ 0  1  0  0  0  0  0 19  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0 20  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0 20  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0 20  0  0  0]
 [ 0  0  0  0  0  0  0  2  0  0  0 18  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 20  0]
 [ 0  0  0  0  0  0  0  1  0  0  0  0  0 19]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0 20]]
                          precision    recall  f1-score   support

  DDoS-ACK_Fragmentation       1.00      0.95      0.97        20
        DDoS-HTTP_Flood        0.83      0.95      0.88        20
        DDoS-ICMP_Flood        1.00      1.00      1.00        20
DDoS-ICMP_Fragmentation        1.00      0.90      0.95        20
      DDoS-PSHACK_Flood        0.95      1.00      0.98        20
       DDoS-RSTFINFlood        1.00      1.00      1.00        20
         DDoS-SYN_Flood        1.00      0.95      0.97        20
         DDoS-SlowLoris        0.83      0.95      0.88        20
DDoS-SynonymousIP_Flood        1.00      1.00      1.00        20
         DDoS-TCP_Flood        1.00      1.00      1.00        20
         DDoS-UDP_Flood        1.00      1.00      1.00        20
  DDoS-UDP_Fragmentation       1.00      0.90      0.95        20
          DoS-SYN_Flood        1.00      1.00      1.00        20
          DoS-TCP_Flood        1.00      0.95      0.97        20
          DoS-UDP_Flood        1.00      1.00      1.00        20

               accuracy                             0.97       300
              macro avg        0.97      0.97      0.97       300
           weighted avg        0.97      0.97      0.97       300
```

Figure 11. Voting Classifier

TABLE I. ANALYSIS OF CLASSIFIERS

Model	ACC (%)	P (%)	R (%)	F1-Score (%)
KNN	96%	96%	96%	96%
SVM	12%	03%	12%	04%
DT	96%	96%	96%	96%
RF	95%	96%	95%	95%
NB	57%	64%	57%	53%
ET	91%	92%	91%	91%
Voting Classifier	97%	97%	97%	97%

V. CONCLUSION

In concluding the evaluation of supervised machine learning models for IoT-DDoS attack prediction, it's evident that certain models excel in accuracy while others showcase strengths in precision, recall, and F1-Score metrics. Notably, K-Nearest Neighbors (KNN), Decision Tree (DT), Random Forest (RF), Extra Trees (ET), and the ensemble Voting Classifier achieved accuracy rates ranging from 91% to 97%, indicating their proficiency in classifying instances accurately. However, when considering precision, recall, and F1-Score, Support Vector Machine (SVM) and Naive Bayes (NB) demonstrate comparatively lower performance. The Voting Classifier stands out as the top-performing model across all metrics, showcasing its reliability and robustness in predicting IoT-DDoS attacks effectively. Its ability to combine insights from multiple classifiers leads to enhanced accuracy, precision, recall, and F1-Score. Conversely, models like SVM and NB, while achieving lower accuracy, show limitations in correctly identifying instances of IoT-DDoS attacks, as reflected in their lower precision, recall, and F1-Score values.

Overall, the evaluation underscores the importance of considering multiple metrics to comprehensively assess the performance of machine learning models in IoT security applications. The Voting Classifier's consistent high scores highlight its potential as a reliable tool for detecting and mitigating IoT-DDoS attacks with precision and efficiency.

VI.REFERENCES

[1] J. Bhayo, S. A. Shah, S. Hameed, A. Ahmed, J. Nasir, and D. Draheim, "Towards a machine learning-based framework for DDOS attack detection in software-defined IoT (SD-IoT) networks," Engineering Applications of Artificial Intelligence, vol. 123, no. July 2022, p. 106432, 2023, doi: 10.1016/j.engappai.2023.106432.

[2] F. S. de Lima Filho, F. A. F. Silveira, A. de Medeiros Brito Junior, G. Vargas-Solar, and L. F. Silveira, "Smart detection: An online approach for DoS/DDoS attack detection using machine learning," Secur. Commun. Netw., vol. 2019, pp. 1–15, 2019.

[3] 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, undefined 2019.

[4] Y. Otoum and A. Nayak, "AS-IDS: Anomaly and signature-based IDS for the internet of things," J. Netw. Syst. Manag., vol. 29, no. 3, 2021.

[5] R. Zagrouba and R. AlHajri, "Machine Learning based attacks detection and countermeasures in IoT," International j. commun. netw. inf. secur., vol. 13, no. 2, 2021.

[6] B. K. Mohanta, D. Jena, U. Satapathy, and S. Patnaik, "Survey on IoT security: Challenges and solution using machine learning, artificial intelligence and blockchain technology," Internet of things, vol. 11, no. 100227, p. 100227, 2020.

[7] S. Pokhrel, R. Abbas, and B. Aryal, "IoT Security: Botnet detection in IoT using Machine learning," arXiv [cs.LG], 2021.

[8] E. Anthi, L. Williams, M. Slowinska, G. Theodorakopoulos, and P. Burnap, "A supervised intrusion detection system for smart home IoT devices," IEEE Internet Things J., vol. 6, no. 5, pp. 9042–9053, 2019.

[9] M. Essaid, D. Kim, S. H. Maeng, S. Park, and H. T. Ju, "A collaborative DDoS mitigation solution based on ethereum smart contract and RNN-LSTM," in 2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS), 2019, pp. 1–6.

[10] U. Javaid, A. K. Siang, M. N. Aman, and B. Sikdar, "Mitigating loT Device based DDoS Attacks using Blockchain," in Proceedings of the 1st Workshop on Cryptocurrencies and Blockchains for Distributed Systems, 2018.

[11] R. Singh, S. Tanwar, and T. P. Sharma, "Utilization of blockchain for mitigating the distributed denial of service attacks," Secur. Priv., vol. 3, no. 3, 2020.

[12] K. Bhardwaj, J. C. Miranda, and A. Gavrilovska, "Towards IoT-DDoS prevention using edge computing," Usenix.org. [Online]. Available: https://www.usenix.org/system/files/conference/hotedge18/hotedge18-papers-bhardwaj.pdf.

[13] P. Kumar, R. Kumar, G. P. Gupta, and R. Tripathi, "A Distributed framework for detecting DDoS attacks in smart contract-based Blockchain-IoT Systems by leveraging Fog computing," Trans. emerg. telecommun. technol., vol. 32, no. 6, 2021.

[14] N.-N. Dao et al., "Securing heterogeneous IoT with intelligent DDoS attack behavior learning," IEEE Syst. J., pp. 1–10, undefined 2021.

[15] V. Adat and B. B. Gupta, "A DDoS attack mitigation framework for internet of things," in 2017 International Conference on Communication and Signal Processing (ICCSP), 2017, pp. 2036–2041.