International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 11 | Nov - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

DDOS Attack Classifier Using Machine Learning

Aryan Chauhan[1], Prince Gandhi[2], Nisarg Patel[3] , Dhara Parikh[4]

[123]*Department of Information Technology, Institute of Information Technology, Krishna School Of Emerging Technology & Applied Research, KPGU University , Varnama, Vadodara, Gujarat, India*

[4] *Assistant Professor, Department of Information Technology and Engineering, Krishna School Of Emerging Technology & Applied Research, KPGU University. Varnama, Vadodara, Gujarat, India*

Abstract - The increasing frequency and complexity of Distributed Denial of Service (DDoS) attacks present substantial challenges to network security. Traditional Intrusion Detection Systems (IDS) often struggle to detect intricate attack patterns in real-time. This research introduces a machine learning-based classifier designed to accurately identify and classify DDoS attacks. Leveraging the Intrusion Detection Evaluation Dataset (CIC-IDS2017), which contains realistic and labeled network traffic, we assess multiple models—Random Forest, Logistic Regression, Gradient Boosting, and Naive Bayes. By incorporating advanced feature selection and hyperparameter tuning, this classifier effectively minimizes false positives and demonstrates strong performance across metrics like accuracy, precision, and recall, making it a promising candidate for real-time DDoS detection in modern networks.

***Key Words*:** Distributed Denial of Service (DDoS), Machine Learning, Intrusion Detection System (IDS), Random Forest, Logistic Regression, Naive Bayes, Gradient Boosting, Exploratory Data Analysis (EDA)

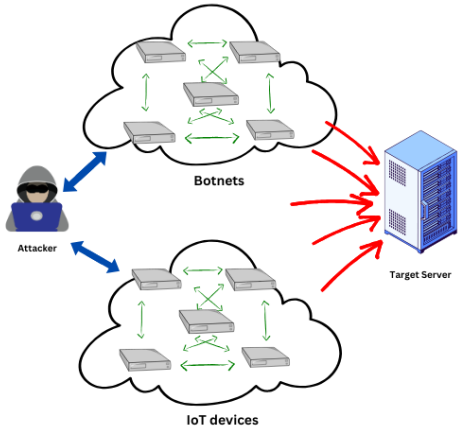
1. Introduction

This Distributed Denial of Service (DDoS) attacks have emerged as a critical challenge in modern network security, aimed at disrupting the availability of services by overwhelming them with excessive, malicious traffic. In a DDoS attack, multiple compromised systems, often organized into a botnet, are used to flood a target—such as a server, network, or application—with an unmanageable volume of requests. This traffic surge causes the target system to slow down or crash, leading to service disruptions that impact businesses, public services, and critical infrastructure. The rise of internet-connected devices, particularly in the Internet of Things (IoT) domain, has exacerbated this issue by providing attackers with a vast pool of unsecured devices that can be easily co-opted into large-scale botnets.

DDoS attacks vary significantly in type and sophistication, generally falling into categories like volumetric attacks, protocol attacks, and application-layer attacks. Volumetric attacks, such as UDP and ICMP floods, aim to exhaust the bandwidth of a network by overwhelming it with massive amounts of data. Protocol attacks exploit vulnerabilities in network protocols, using methods like SYN floods and fragmented packet attacks to consume server resources. Application-layer attacks, such as HTTP floods, target specific functions of applications, causing a slower but equally disruptive denial of service. The complexity and variety of these attack types make DDoS attacks particularly challenging to detect and mitigate, as they require systems that can handle a wide range of patterns and adapt to evolving tactics.

Intrusion Detection Systems (IDS) and Intrusion Prevention Systems (IPS) are fundamental defenses against DDoS attacks. However, traditional IDS/IPS approaches often rely on signature-based methods, which detect attacks based on known patterns or rules. These approaches are limited in their ability to recognize new, emerging attack patterns and are prone to high false positive rates in complex network environments. To address these limitations, security researchers are increasingly turning to machine learning-based methods, which use data analysis and pattern recognition to detect anomalous behavior that may signify a DDoS attack. Unlike rule-based methods, machine learning models can adapt to new attack patterns, making them well-suited for the dynamic and evolving nature of network traffic.

The goal of this study is to leverage machine learning techniques to classify DDoS attacks, offering a more flexible and robust approach to network defense. Machine learning classifiers analyze network traffic data, learning patterns associated with malicious activity, and can thereby achieve higher detection accuracy than traditional techniques. By employing algorithms such as Random Forest, Gradient Boosting, Logistic Regression, and Naive Bayes, this study seeks to identify classifiers that not only detect DDoS attacks with high accuracy but also minimize false positives, which is critical for reducing unnecessary alerts and resource consumption.



The architecture of a DDoS attack typically includes multiple phases: the recruitment of devices into a botnet, the orchestration of the attack, and the delivery of coordinated traffic to the target system. The following diagram illustrates the architecture of a DDoS attack, showing how compromised devices (“IoT devices”or "bots") are directed by an attacker to simultaneously flood a target. This multi-layered structure, which often involves traffic from various IP addresses, complicates detection efforts, as it can make malicious traffic appear similar to legitimate user activity.

With the continuous rise in DDoS incidents across sectors, from finance to healthcare, it is essential to develop classifiers that can offer real-time, accurate detection. By exploring the performance of different machine learning models, this research aims to contribute insights that can guide the development of more effective and adaptive DDoS mitigation strategies, ultimately helping to safeguard critical network infrastructures against future attacks.

2. Related Work

Several studies have highlighted the importance of robust DDoS detection mechanisms, increasingly favoring machine learning techniques over traditional rule-based or statistical methods. These conventional approaches, reliant on predefined signatures and thresholds, struggle to adapt to the evolving nature of DDoS attacks, often resulting in high false-positive rates. In contrast, machine learning methods offer greater adaptability and precision by analyzing complex traffic patterns, enabling more effective detection with minimal human intervention.

Recent research has applied various machine learning classifiers to DDoS detection, demonstrating improved accuracy. [1] utilized models like Random Forest and Support Vector Machines to enhance detection rates, while [2] showed that machine learning models can achieve high accuracy for medium-scale attacks. However, scalability remains a challenge in large-scale attack scenarios, highlighting the need for more efficient, adaptable classifiers.

[3] researchers found that supervised machine learning models outperform rule-based methods in IoT environments. However, they noted a reliance on simulation-based validation, which limits generalizability to real-world conditions. This underscores the need for models that handle diverse attack data effectively, a challenge our study addresses through feature selection and preprocessing techniques to improve model generalization.

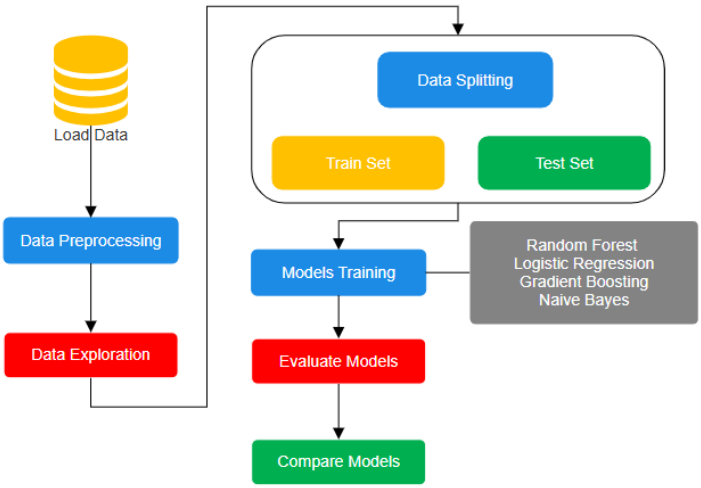
Feature selection has been emphasized in studies such as [4], where techniques like PCA and SVD optimized computational efficiency without sacrificing accuracy. Similarly, our approach integrates feature encoding, normalization, and high-impact feature selection, reducing complexity and enabling models like Gradient Boosting to achieve high precision and recall, as confirmed by our findings.

While deep learning models, such as Convolutional Neural Networks (CNN), have shown promise in DDoS detection, as noted in [5], they require extensive computational resources, posing challenges for real-time applications. To address this, our study focuses on efficient machine learning models like Random Forest and Gradient Boosting, which balance accuracy and processing demands, making them suitable for real-time and scalable DDoS detection

In summary, while existing work has advanced DDoS detection, gaps remain in achieving scalability, real-world generalizability, and computational efficiency. Our solution addresses these challenges through a carefully selected ensemble of machine learning models optimized by thorough preprocessing and feature selection. These improvements lead to a high-accuracy classifier that meets the demands of real-time DDoS detection in diverse network environments, reducing false positives and enhancing adaptability across attack scenarios.

3. Methodology

In this research, we designed a machine learning-based classifier to detect and classify DDoS attacks by following a structured approach involving dataset loading, data preprocessing, model training, evaluation, and comparison.



A. Dataset

For this research, we utilized the [6] Friday-WorkingHours-Afternoon-DDos dataset from the CICIDS2017 dataset collection by the Canadian Institute for Cybersecurity (CIC). Widely adopted for its realistic simulation of modern network conditions, the CICIDS2017 dataset is commonly used in network security research and encompasses various types of attacks, including DDoS. Specifically, this subset consists of labeled network traffic data collected during simulated DDoS attack scenarios.

The dataset contains 79 features capturing detailed network parameters such as packet length, flow duration, flag counts, and data rates, which serve as inputs for machine learning models to distinguish between benign and malicious traffic. Key features include Flow Duration, Total Fwd Packets, Flow IAT Mean, and Packet Length Mean, among others, which provide crucial insights into traffic patterns typical of DDoS attacks.

This dataset was selected for its thorough labeling and comprehensive feature set, enabling robust model training and evaluation for DDoS attack detection.

B. Data Preprocessing

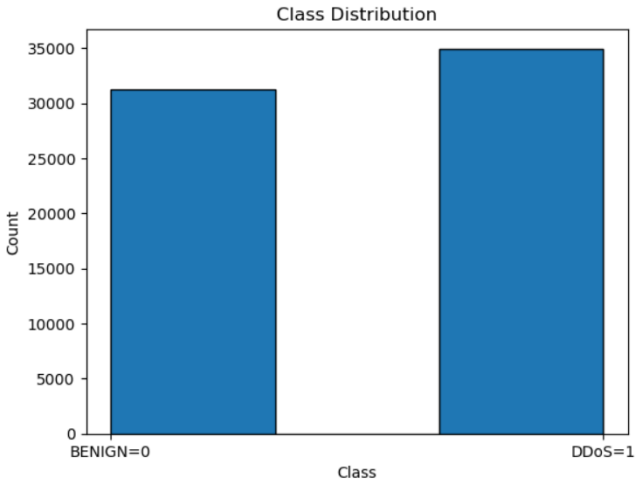
Data preprocessing is a crucial step to prepare the dataset for model training. This process involves cleaning the data, encoding labels, normalization, data exploration, and splitting the dataset, as described below:

Column Cleaning: Leading and trailing spaces in column names were removed to standardize feature labels and avoid errors during data manipulation.

Target Verification: The target classes (Label column) were inspected to confirm unique values, showing 'BENIGN', 'DDoS', and NaN as possible values. Missing values were subsequently removed to ensure a clean dataset.

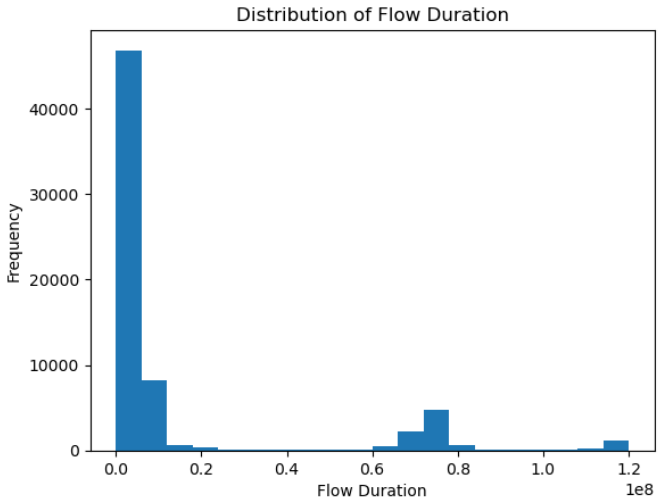
Data Cleaning: Missing values were dropped from the dataset, resulting in a clean dataset ready for analysis. Non-numeric columns, such as Label, were converted into numerical values (BENIGN = 0, DDoS = 1) for model compatibility.

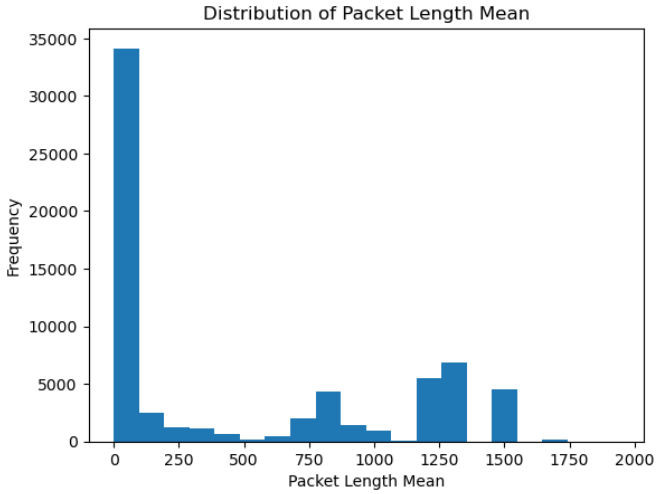
Class Distribution: The distribution of classes in the dataset was checked post-cleaning to ensure an adequate representation of each class for training purposes.

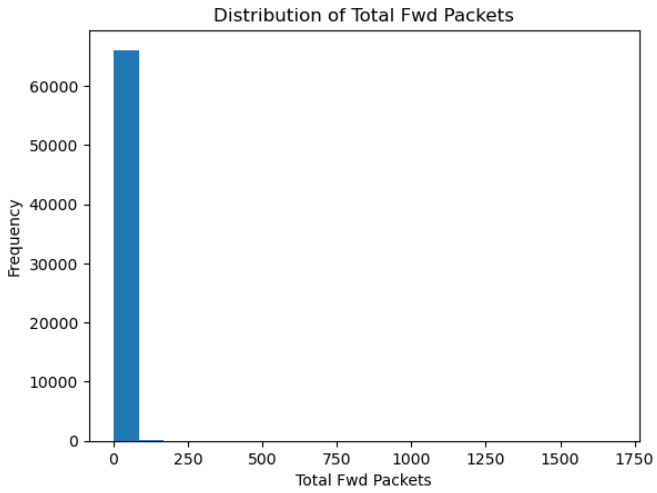


Exploratory Data Analysis (EDA): EDA was performed to understand the data distribution and statistical properties of each feature. Descriptive statistics were generated to provide a summary of each feature's distribution, aiding in identifying relevant features for detecting DDoS patterns.

Distribution histograms for key features, such as Flow Duration, Total Fwd Packets, and Packet Length Mean, illustrate the spread and concentration of values, helping to visualize patterns typical of benign and DDoS traffic. These three features are highlighted as examples of primary indicators for DDoS detection; however, similar analyses were conducted for all major features to provide a comprehensive understanding of the dataset.







.

Data Splitting: The data was split into training and testing sets using a **70:30 ratio**, with the training set used to fit models and the test set reserved for evaluating model performance. Features (X) and target (y) were separated before applying the split, ensuring clear boundaries between predictors and the label.

C. Model Training

To classify DDoS attacks effectively, we trained four machine learning models, each with unique strengths in handling binary classification tasks. These models were selected for their diversity in approach, allowing us to compare their effectiveness in distinguishing between benign and malicious traffic. Each model was trained on the preprocessed training data, and hyperparameter tuning was applied to maximize accuracy and generalizability.

Random Forest: An ensemble method that builds multiple decision trees and averages their predictions to improve accuracy and reduce overfitting.

Logistic Regression: A linear model for binary classification, predicting class probability with a logistic function.

Gradient Boosting: An iterative method that adds models sequentially to correct previous errors, minimizing a loss function.

Naive Bayes: A probabilistic model that uses Bayes' theorem, assuming feature independence.

D. Model Evaluation and Comparison

To assess model performance, several evaluation metrics were employed to capture various aspects of classification quality:

Accuracy: Measures the proportion of correct predictions out of total predictions.

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Samples}}$$

Precision and Recall: These metrics evaluate the classifier's ability to identify DDoS attacks accurately.

Precision (positive predictive value) represents the relevance of positive predictions, defined as

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

Recall (sensitivity) measures how effectively the model detects DDoS attacks

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

F1 Score: The harmonic mean of precision and recall, providing a single balanced metric, especially useful for imbalanced datasets.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

ROC-AUC Curve: The Receiver Operating Characteristic (ROC) curve plots the true positive rate against the false positive rate, with the Area Under the Curve (AUC) providing a single score to evaluate overall performance across various classification thresholds.

Confusion Matrix: A confusion matrix is a table that visually represents the classification performance by summarizing true positives, true negatives, false positives, and false negatives. It provides insights into each model's strengths and weaknesses, showing how well the classifier distinguishes between DDoS and non-DDoS traffic.

True Positives (TP): Correctly identified DDoS attacks.
True Negatives (TN): Correctly identified benign traffic.
False Positives (FP): Benign traffic incorrectly classified as DDoS.
False Negatives (FN): DDoS traffic incorrectly classified as benign.

Each model was evaluated based on the metrics mentioned above, and results were visualized using ROC-AUC curves and confusion matrices. The Random Forest and Gradient Boosting models outperformed Logistic Regression and Naive Bayes, achieving high precision, recall, and AUC scores, indicating they are the most suitable for real-time DDoS detection applications.

.

4. Experiment Results

The performance of each model was evaluated using several key metrics: accuracy, precision, recall, F1 score, and ROC-AUC, which collectively provide insights into the models' abilities to accurately classify DDoS and benign traffic. Below is a detailed analysis of each model's performance and a comparison of their strengths and weaknesses for DDoS detection.

Naive Bayes achieved high recall, effectively identifying most DDoS traffic, but had lower precision, leading to more false positives. This is likely due to its assumption of feature independence, which can be limiting in complex data.

Logistic Regression showed balanced performance with high recall and good precision, making it a solid baseline for DDoS detection, though it lags behind ensemble models. It's efficient and suitable for low-resource environments.

Gradient Boosting excelled with near-perfect results, offering 100% precision and very high recall, minimizing false positives and negatives. Its iterative approach makes it ideal for high-accuracy scenarios.

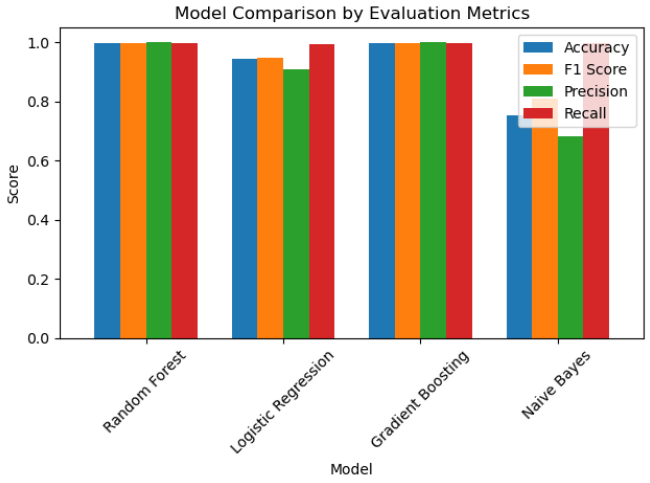
Random Forest's performance closely matched Gradient Boosting, with high precision and recall, making it highly reliable for DDoS detection. Its ensemble of decision trees prevents overfitting and captures DDoS traffic patterns well.

Comparative Analysis :

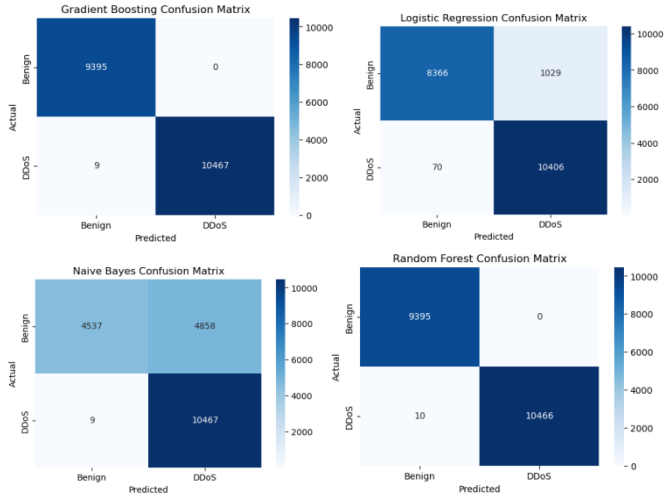
The following table summarizes the accuracy, precision, recall, and F1 score for each model:

Model	Accuracy	Precision	Recall	F1 Score
Naive Bayes	75.51%	68.30%	99.91%	0.8114
Logistic Regression	94.47%	91.00%	99.33%	0.9498
Gradient Boosting	99.95%	100%	99.91%	0.9996
Random Forest	99.95%	100%	99.90%	0.9995

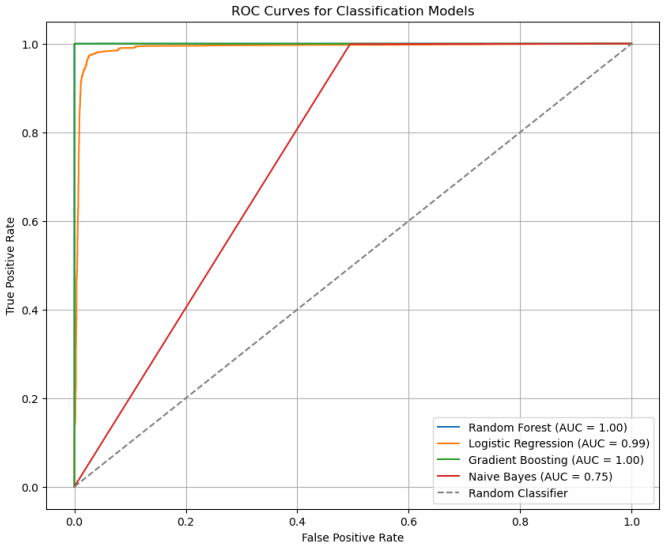
Both Gradient Boosting and Random Forest achieved the highest performance in terms of accuracy and F1 score, with minimal false positives, making them well-suited for precise and reliable DDoS detection. Naive Bayes, while having a high recall, was prone to false positives, whereas Logistic Regression provided a solid balance across metrics.



Confusion Matrices:



ROC-AUC Curve:



The evaluation metrics and visual comparisons indicate that Gradient Boosting and Random Forest are the most effective models for DDoS attack classification, achieving the highest accuracy, precision, and F1 scores. These models demonstrate minimal false positives and false negatives, making them reliable candidates for real-time DDoS detection systems. Naive Bayes and Logistic Regression, while performing adequately, were outperformed by the ensemble models, particularly in handling the complexity of DDoS traffic patterns.

5. Conclusion

In conclusion, this study successfully developed and evaluated machine learning classifiers to detect DDoS attacks using the CICIDS2017 dataset, focusing on Random Forest and Gradient Boosting as top-performing models. These classifiers demonstrated strong accuracy, precision, and recall, with minimal false positives, making them suitable for real-time DDoS detection. By addressing the challenges of high accuracy and low false positive rates, this approach shows promise in maintaining the integrity and availability of network services. Future work could explore deep learning models and hybrid methods to improve adaptability and generalization across complex datasets, as well as optimization algorithms to enhance computational efficiency. Testing these models in live network environments will be critical to ensure their practical applicability and robustness under real-world conditions

References

1) DDoS Attacks Detection and Classification Based on Deep Learning Model - Research Square, Tlemcen University, 2023
2) An Approach of DDoS Attack Detection Using Classifiers - Emerging Research in Computing, Information, Communication, and Applications, Springer India , 2015
3) Survey and classification of Dos and DDos attack detection and validation approaches for IoT environments - Elsevier , 2024
4) Denial of Service Attack Classification Using Machine Learning with Multi-Features - MDPI , 2022
5) Classification of DDoS attack traffic on SDN network environment using deep learning - Springer Open, 2024
6) Dataset : https://www.unb.ca/cic/datasets/ids-2017.html