# **DDoS Attack Detection and Classification**

## Abstract

This report presents a comprehensive project in the field of machine learning, centred around the analysis and mitigation of Distributed Denial of Service (DDoS) attacks. The project starts with the exploration and visualization of a provided DDoS dataset, aiming to gain insights into the characteristics and patterns of the network traffic associated with such attacks. The visualization and exploration phase leverages data visualization techniques to unveil key features and trends within the dataset.

Subsequently, the project delves into the application of supervised learning techniques for classification purposes. Multiple machine learning models are employed to classify network traffic instances: KNN, Support Vector Machines, Gaussian Naive Bayes and Random Forests. The evaluation of these models involves metrics such as accuracy, precision, recall, and F1 score, providing a comprehensive assessment of their performance in identifying DDoS attacks.

Furthermore, the project incorporates clustering techniques to uncover hidden structures within the dataset. By applying clustering algorithms such as K-means, DBSCAN and Gaussian Mixture; the project aims to group similar network traffic instances together. Evaluation of clustering results is performed using appropriate unsupervised and supervised metrics.

## Section 1

## Section 2

In this project, to solve the multilabel classification problem, four supervised ML classifiers have been evaluated, which are K-Nearest-Neighbours-Classifier (KNN), Random Forest Classifier (RF), Support Vector Machine (SVC) and Gaussian-Naive-Bayes (GNB). The following are the metric used for model evaluation:

* Accuracy:
* Precision:
* Recall:
* F1-score:

Firstly, the dataset (pca\_dataframe.csv) is split in training and test set and they are used to train and evaluate the models (default parameters) respectively.

Immagine che contiene testo, schermata, Parallelo, numero

Descrizione generata automaticamenteImmagine che contiene testo, schermata, Parallelo, numero

Descrizione generata automaticamente

KNN, RFC and SVC have similar results on the test set with KNN and RFC that reach a score above 80% for all the metrics evaluated. GNB, since it starts from the assumption of Gaussian distribution of data and related probability independence among features, is the one with worst performance with an average score of 60% for the metrics involved in the evaluation.

Considering the confusion matrix on test set, it is possible to visually highlight which class of traffic are misclassified.

Immagine che contiene testo, schermata, quadrato, diagramma

Descrizione generata automaticamenteImmagine che contiene testo, schermata, quadrato, diagramma

Descrizione generata automaticamente

Immagine che contiene testo, schermata, quadrato, diagramma

Descrizione generata automaticamenteImmagine che contiene testo, schermata, quadrato, diagramma

Descrizione generata automaticamente

There are few features that are classified erroneously (<90% of correct predictions) in all the models:

* ddos\_dns (true) misclassified with: ddos\_mssql, ddos\_netbios, ddos\_ntp, (only in SVC ddos\_snmp), ddos\_ssdp
* ddos\_ldap (true) misclassified with: ddos\_mssql, ddos\_ssdp
* ddos\_mssql (true) misclassified with: ddos\_dns (only in KNN), ddos\_ldap, ddos\_ssdp
* ddos\_ssdp (true) misclassified with: ddos\_mssql, ddos\_ldap, ddos\_dns (not in SVC)
* ddos\_udp (true): misclassified with: ddos\_udp\_lag (>40% erroneous classifications), ddos\_netbios (not in SVC)
* ddos\_udp\_lag (true): misclassified with: ddos\_udp\_lag (>40% erroneous classifications), ddos\_netbios (not in SVC)

ddos\_udp and ddos\_udp\_lag are the ones that are misclassified the most among each other. It is understandable from the nature of this kind of flows which are strictly related. The first one is an actual DDoS attack that exploits the vulnerabilities of UDP protocol, while the other (UDP DDoS lag) a type of DoS attack that floods a target server with UDP packets with an invalid checksum. This can cause the server to spend time processing the invalid packets, which can slow down or even crash the server.

From this point on, the GNB model isn’t considered anymore due to its low performance in respect of the others.

After the first model evaluation, we proceeded with hyperparameters tuning for each model to try to increase their performances on our DDoS dataset. We choose to use a Grid Search algorithm that takes care of performing cross validation trying to reach more reliable performance estimates, reduce overfitting, and contributing to a better understanding of a model's generalization capabilities. The following are the hyperparameters tuned in the process which involved sklearn.GridSearchCV():

* KNN:
  + “n\_neighbours”: [3, 5, 7]
  + “weights”: [“uniform”, “distance”]
  + “p”: [1, 2]
* RFC:
  + “criterion”: [gini, entropy]
  + “n\_estimators”: [50, 100]
  + “max\_depth”: [None, 10]
  + “min\_samples\_split”: [2, 3]
* SVC:
  + “C”: [0.1, 1, 10]
  + “kernel”: [rbf, poly]

For each model, as a result, the best set of hyperparameters are:

* KNN:
  + “n\_neighbours” = 7
  + “p” = 1
  + “weights” = “uniform”
* RFC:
  + “criterion” = “entropy”
  + “max\_depth” = None
  + “min\_samples\_split” = 3
  + “n\_estimators” = 100
* SVC:
  + “C” = 10
  + “kernel” = “rbf”

Immagine che contiene testo, schermata, numero, Parallelo

Descrizione generata automaticamenteImmagine che contiene testo, schermata, Parallelo, numero

Descrizione generata automaticamenteThe models have been initialized with their best hyperparameters and evaluating them on the dataset has produced the following results:

As a result, there is little or no significant improvement for all the three models in the test set predictions. This kind of behaviour can be caused from the complex dataset nature, so models are not able to achieve a higher level of performance on the evaluated metrics.