Classifying Action of Internet Firewall Using Machine Learning Models

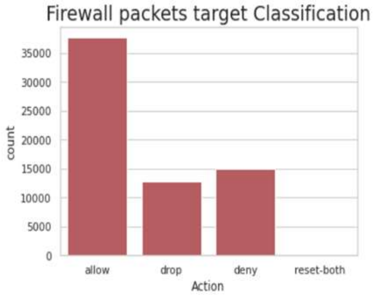
Geetanjali Kakda, Dr. Eman Abdelfattah
School of Computer Science & Engineering
Sacred Heart University, Fairfield, CT

Abstract

A firewall is a type of security mostly located at the entry and exit points of a network. This is necessary to improve the security of the network and defend against cyber threats. This study aims to classify the internet traffic using different Machine Learning classification models into four categories based on 11 features. The models applied are – K-Nearest Neighbor, SGD, Decision Tree, Random Forest , XGBoost, Support Vector Machine. The XGBoost model achieved the highest accuracy, recall, precision,F1score. Decision Tree has the least time complexity.

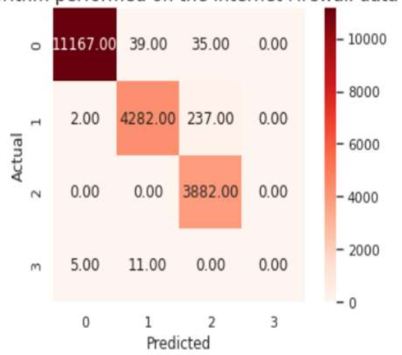
Dataset Description

- This data set was collected from the internet traffic records on a University's firewall.
- This is a multi-class data set.
- It consists of 65532 instances.
- There are 12 features in total.
- Action feature is used as a class.



Experimental Results & Analysis:

Confusion matrix for the SGD algorithm performed on the Internet Firewall data



Confusion matrix for the XgBoost algorithm performed on the Internet Firewall data

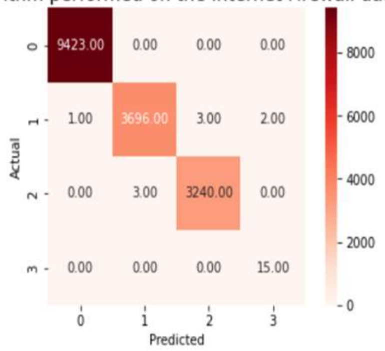
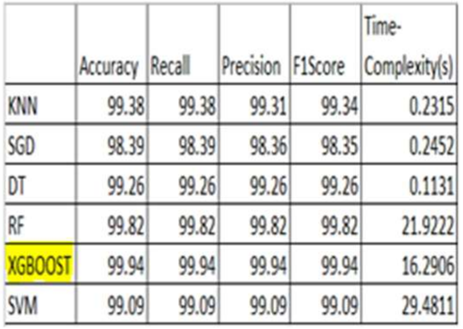
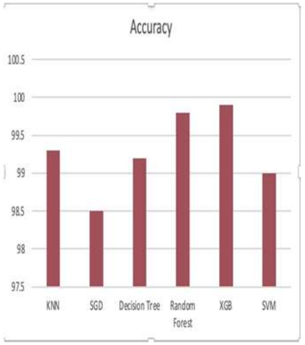


TABLE SHOWS VALUES OF ACCURACY, RECALL,PRECISION,F1SCORE AND TC



	Accuracy	Recall	Precision	F1Score	Time-Complexity(s)
KNN	99.38	99.38	99.31	99.34	0.2315
SGD	98.39	98.39	98.36	98.35	0.2452
DT	99.26	99.26	99.26	99.26	0.1131
RF	99.82	99.82	99.82	99.82	21.9222
XGBOOST	99.94	99.94	99.94	99.94	16.2906
SVM	99.09	99.09	99.09	99.09	29.4811

BAR CHART FOR ACCURACY KNN, SGD, DT , RF , XGBOOST and SVM MODELS



Conclusion:

In conclusion the study shows that the both Random Forest and XGBoost Models perform better than other models. XGBoost performs has the best performance in terms of accuracy of 99.9450 %. KNN performs best in terms of runtime. In future work larger data can be handled extracted from different firewalls to achieve high performance of classification.

References:

[1] R. F. Naryanto and M. K. Delimayanti, "Machine Learning Technique for Classification of Internet Firewall Data Using RapidMiner, " 2022 6th International Conference on Electrical, Telecommunication and Computer Engineering (ELTICOM), Medan, Indonesia, 2022, pp. 155-159,doi: 10.1109/ELTICOM57747.2022.10037798.C