Proceedings of the 5th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/29/20230972

Machine learning-based DDoS detection for IoT networks

Yafei Xie

Department of Computer Science, University College London, London, WC1E 6BT, UK

oliver.xie.22@ucl.ac.uk

Abstract. DDoS attacks are one of the most dangerous threats to IoT networks, and they involve using attacker-controlled botnets to flood the network with malicious traffic that denies legitimate services. The global DDoS landscape is rapidly evolving, and it has become increasingly important for devices to quickly identify the types of DDoS attacks they face so that they can choose and implement effective countermeasures against known attacks. Machine learning has emerged as a popular approach for detecting DDoS traffic in IoT networks. This paper implements four machine learning models, namely Support Vector Machine (SVM), Decision Tree, Long Short-Term Memory (LSTM), and Random Forest, to perform multiclass classification for DDoS attack detection. The study uses the CICDDoS2019 dataset for evaluation. The results show that all four models can detect most types of DDoS traffic effectively. The Random Forest model achieves the highest overall accuracy of 99.32%, followed by the Decision Tree model with an accuracy of 99.10%. The LSTM and SVM models have slightly lower accuracies at 98.20% and 93.00%, respectively. The study also evaluates the models' performance in terms of precision, recall, and F1 score. Decision Tree outperforms the other models in precision, while Random Forest has the highest recall score. Moreover, the Random Forest model performs the best in terms of the F1 score. In conclusion, this paper demonstrates the effectiveness of machine learning-based approaches for DDoS detection in IoT networks using four popular models. The results illustrate the potential for these models to provide reliable and accurate detection of DDoS traffic, thus enabling effective countermeasures to be taken against this type of attack.

Keywords: DDos, classifier, SVM, decision tree, LSTM, random forest.

1. Introduction

The statistical data about the amount of Internet of Thing (IoT) devices has shown an increasing trend in recent years. In 2030, the amount of Internet IoT will be expected to reach 25 billion [1]. However, due to the limited resource in computation and storage, it is not practical to apply the complex defence system in IoT devices for security purpose [2]. Therefore, the IoT's increasing quantity and its resource limitation turned IoT devices into popular targets for adversaries to attack. As a result, the vulnerable IoT device can become a weakness point in the security of a whole network system.

Distributed Denial of Service attack (DDos) is one of most serious attack about the IoT devices that aims to break the availability of the devices. The attacker would use a huge botnet to perform flooding attack which deny the legitimate user to access the deserved service. One well-known example is Mirai botnet [3]. It can launch massive DDoS attacks to infect vulnerable IoT devices that have weak or default

DOI: 10.54254/2755-2721/29/20230972

usernames and passwords. Moreover, once the IoT device is infected, this device would become part of the botnet which can perform powerful DDos attacks.

Currently, the volume of global DDoS attack has rapidly increased. The quantity of traffic resulting from DDoS attacks rose by 79% overall [4]. The network security report of Microsoft in 2022 showed that Microsoft defended against an average of 1,435 DDos threats every single day [5].

There are two categories of DDos attacks, which are reflected based and exploitation based [6]. The main difference between the two types of DDos is the method of execution. In reflected-based, the attacker requests huge amount of servers and uses massive responses to flood the victim, which is called reflection amplification. However, exploitation based DDos directly exploits the vulnerabilities in the victim's device. Additionally, DDos in these two categories can be divided into many different types depending on the protocols used. Exploitation Based has SYN flood, UDP flood, UDP-lag. Reflected based DDos involves DNS, MSSQL, SSDP, LDAP, NTP, TFTP and SNMP.

Detecting the abnormal traffic of various DDos attacks is a crucial step in preventing DDoS attacks. Nowadays, machine learning is one popular approach used in IoT network [7]. The adaptability and scalability are the major advantages of using machine learning. Algorithms that utilise machine learning can adapt to new attack patterns and learn from previous attacks, enabling them to detect and mitigate future attacks more effectively. This ensures that network defences remain effective and up-to-date. If the devices can rapidly determine the type of DDoS attack they encountered, they can select the most appropriate countermeasure, which is more effective than general solution. Therefore, this paper applied four algorithms, SVM, Decision Tree, Random Forest and LSTM to perform multiclass classification for detecting different types of DDos attack, and compared the results from four algorithms.

2. Related work

Kolias et al. presented the example of DDos attack in IoT network using Mirai botnet. Their work showed the damage caused by Mirai botnet, the processes of implement the DDos attack using Mirai bonet and other relevant botnets, which made people aware of the danger of botnets and called for safety in IoT network.

Vishwakarma and Jain summarised several well-known varieties of DDos attacks based on the protocols, which are SYN, DNS, NTP, UDP, HTTP and ICMP [8]. This work also summarised the current countermeasures in 2019 and compared machine learning approaches with other countermeasures. The results showed that it was popular trend for using machine learning countermeasures against the DDos attack in IoT devices. In 2023, Kumari and Jain did a more comprehensive study in this aspect. Their work provided more details about the DDos attack and more defence mechanisms.

Suresh and Anitha selected 23 features from the DDos traffic data using chi-square and Information gain and applied 6 types of models to detect the DDos attacks. In their results, Fuzzy c-means method had the best accuracy which was 98.7%. However, this work was binary classification, it could not detect which type of DDos attack was implemented [9].

In 2018, Doshi et al. used 5 different algorithms to implement the DDos traffic detection in IoT network. One important strength in this paper was that authors limited the memory and used lightweight features selected from self-collected traffic data. In their results, 4 models reached the 99% accuracy. The classification was also binary [10].

Sharafaldin et al. produced a new DDos traffic dataset which mitigated the weakness of existing datasets in 2019. This dataset involved 11 types of DDos traffic and 80 features. Many research used this data set to train the machine learning models.

Tuan et al. implemented the DDos detection using SVM, Naïve Bayes, decision tree, neural network and unsupervised learning [11]. They work involved two datasets, KDD99 and UNBS-NB 15 datasets. From the results, their study showed that the unsupervised learning could perform better in several aspects. Moreover, the results also demonstrated KDD99 dataset had better performance [12].

Chen et al. implemented a DDos detection system for IoT devices with multiple layers. The first layer was an authentication mechanism to check protocols, the second layer was DDos traffic detection

DOI: 10.54254/2755-2721/29/20230972

using decision tree algorithm, the last layer is the traffic blocking rules in the SDN controller. The classification was binary, and the accuracy scores could reach 97% for sensor data flooding and 99% for network data flooding. Implementing a blocking mechanism after detection is one important strength in this study.

Gaur and Kumar analysed the machine learning classifiers in two aspects using CICDDoS2019 dataset [13]. Firstly, they used 4 algorithms to implement the DDos detection which ware KNN, eXtreme gradient boosting, decision tree and random forest. Secondly, they used three approaches to choose features which were Chi-square, ExtraTreeClassifier and Analysis of variance. These three selection algorithms were used with each machine learning classifier to select 20 features. Compared with results, eXtreme gradient boosting combined with Analysis of variance had the highest accuracy (98%) and F1 score (99%).

3. Methodology

3.1. Dataset

The data used for training and testing purpose is from CICDDoS2019 [6]. One major advantage for this dataset is that it covers both categories of DDos attacks and contains 11 types of DDos traffic data. In order to perform multiclass classification, this study chose 10 types of DDos attack data, which were DNS, UDP, NTP, SNMP, UDP-lag, MSSQL, SYN, LDAP, NetBIOS, SSDP. There is a few normal traffic labelled "BENIGN" in each type of attack data, and a very small amount of WebDDos attack data in UDP-lag traffic. Therefore, there are 12 labels in the dataset. Due to huge amount of traffic data, random selection function was used to pick the similar size of traffic data in each type of DDos traffic file to reduce the computational cost. Then these selected data were combined into a new dataset. After that, the new dataset was split randomly, 70% of the data is training set and 30% is testing set.

3.2. Feature selection

In the original traffic data, there were approximately 87 features, and it was not practical to consider all the features in the training. To reduce the computational cost and complexity, a build-in function from scikit-learning packets, ExtraTreeClassifier, was used to select the best 20 features [14]. This method works by constructing a huge number of decision trees and selecting the features that are most frequently used to divide the trees. The best 20 features are shown in figure 1 with their standard importances.

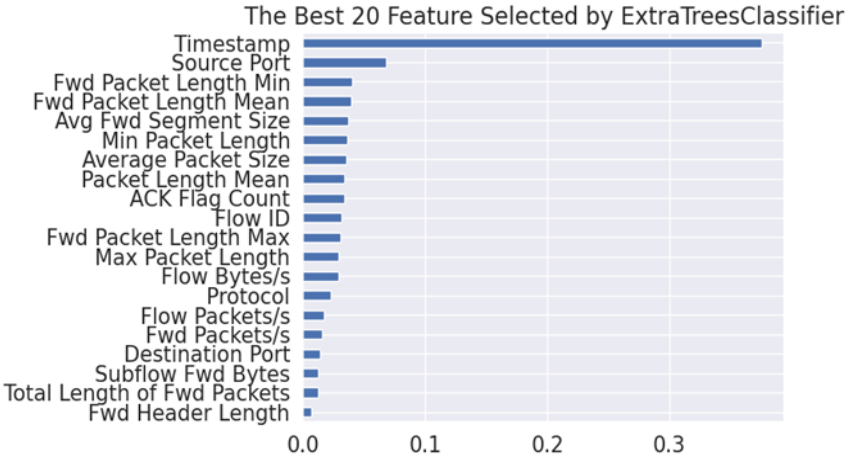

Figure 1. Best 20 Features selected.

3.3. Machine learning algorithms

This section will demonstrate four machine learning algorithms that implemented DDos detection on different types of traffic. This study was implemented in a virtual machine environment, Google

DOI: 10.54254/2755-2721/29/20230972

Collaboratory, with 50 G RAM and 225 G storage [15]. Relevant Python packets were applied to achieve following models.

3.3.1. Support vector machines. Support Vector Machine can identify the hyperplane that maximises this distance while correctly classifying the different type of DDos traffic data. It can handle large datasets efficiently and are less prone to overfitting compared with other models. To implement this model, this study used the LinearSVC function from the scikit-learn package [14]. LinearSVC is a linear SVM algorithm, and its parameters were adjusted to prevent overfitting and achieve optimal performance on the DDoS dataset. In this algorithm, the loss function was square-hinge. In order to perform multiclass classification, the multiclass parameter was set to "ovr" (one-vs-rest).

3.3.2. Decision tree. Decision trees are supervised learning algorithms which use tree-like models of decisions and their possible outcomes to create a predictive model. The labels of different DDos traffic formed the leaf nodes in the decision tree, and each branch showed the consequence of the test. This work chose the function, DecisionTreeClassifier, in scikit-learn package to implement the decision tree.

3.3.3. Random forest. Random Forest is the aggregation of decision trees which are selected random from the training set and trained on a random subset of features. In this aggregation, each tree produced one consequence which predicted the possible type of DDos traffic. Then the final predicted label derived from the majority vote of all trees in the Random Forest. This algorithm can hand the large data set which contains outlier and missing values. Random Forest algorithm was deployed using the RandomForestClassifier function from the scikit-learn package [14].

3.3.4. Long short term memory. Long Short-Term Memory (LSTM) is a type of neural network that can mitigate the issue of vanishing and exploding gradient happened in recurrent neural networks. LSTM contains input, forget and output gates that can control the adding, discarding and outputting operations on the data.

In this study, the LSTM algorithm was implemented using the Keras library from TensorFlow. In LSTM model, eight units were used in each layer, and the Dropout function were applied after the layers to prevent the overfit. To implement the multiclass classification, the softmax was used as the output layer. This model used cross-entropy as the loss function and adam as an optimizer.

3.4. Evaluation

To evaluate the performance of three algorithms, this study used two approaches. The first method was calculating the accuracy score for each model. Function “accuracy_score” from the scikit-learn package to compute the score [14]. It takes the true label from the testing set and predicted label after training as inputs, and outputs the closeness between these two labels. If the all the predicted labels equal to the true labels in the testing set, the accuracy score is 1 otherwise score is 0.

The second approach is the F1 score, which is computed from precision and recall values. The precision of the model may be measured by calculating the percentage of true positives among all of the model's positive predictions. The recall is the percentage of true positive samples in the dataset that include genuine positives relative to the total number of actual positive samples. The F1 score ranges from 0 to 1, with greater scores denoting superior performance. A flawless F1 score of 1.0 demonstrates that the model's precision and recall are perfect, whereas a score of 0 shows that the model's output is entirely inaccurate.

4. Results and analysis

This section will demonstrate and analyse the results from different machine learning models. As mentioned previously, there was a few traffic labelled WebDDos which only presented in the data file of UDP-lag. The amount of WebDDos traffic in training set is very small. Models did not have enough

DOI: 10.54254/2755-2721/29/20230972

data for training to detect the WebDDos traffic. Therefore, this study did not care about the detection result on WebDDos traffic.

4.1. SVM

The accuracy score of Support Vector Machine is 93.002%, which is the lowest score compares with other models. The table 1 below shows the evaluation results on SVM for each type of the DDoS attacks in decreasing order. The F1-Score for SYN attack is 1.00 which means the SVM can perfectly detect the SYN flooding on the system. Because the F1-Score of BENIGN is 0.92, SVM algorithm can have good performance on detection between normal and abnormal traffic. The figure 2 shows the confusion matrix of prediction results by SVM on a small part of training data.

Table 1. Evaluation results of SVM.

	Precision	Recall	F1 Score
Syn	1.00	1.00	1.00
NTP	1.00	0.98	0.99
SNMP	0.97	0.95	0.96
LDAP	0.95	0.93	0.94
NetBIOS	0.88	0.99	0.93
UDP-lag	0.99	0.85	0.92
BENIGN	0.96	0.90	0.92
DNS	0.91	0.91	0.91
MSSQL	0.88	0.92	0.90
UDP	0.81	0.96	0.88
SSDP	0.95	0.80	0.87
WebDDos	0.48	0.57	0.52

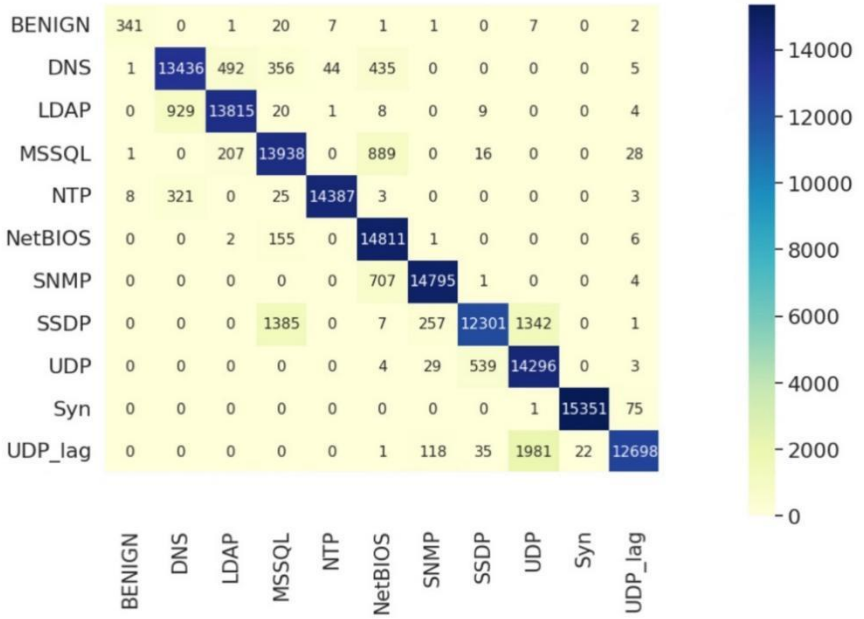

Figure 2. Confusion matrix of prediction results from SVM.

4.2. Decision tree

The accuracy score of decision tree is 99.099%, which is higher than SVM. The table 2 below shows the evaluation results on decision tree for each type of the DDoS attacks in decreasing order. The most

DOI: 10.54254/2755-2721/29/20230972

of F1-Scores were really high except BENIGN. These results means that this model can well detect different types of DDos attack, but it cannot distinguish the normal traffic properly. One important reason is that the amount of normal traffic was smaller than other attack traffic. The figure 3 shows the confusion matrix of prediction results by decision tree on a small part of training data.

Table 2. Evaluation results of decision tree.

	Precision	Recall	F1 Score
Syn	1.00	1.00	1.00
NTP	0.98	1.00	0.99
SNMP	0.99	0.99	0.99
LDAP	0.99	0.99	0.99
NetBIOS	0.99	0.99	0.99
UDP-lag	0.99	0.99	0.99
SSDP	0.99	0.99	0.99
DNS	0.99	0.99	0.99
MSSQL	0.99	0.99	0.99
UDP	0.99	0.99	0.99
BENIGN	1.00	0.43	0.60
WebDDos	0.67	0.29	0.40

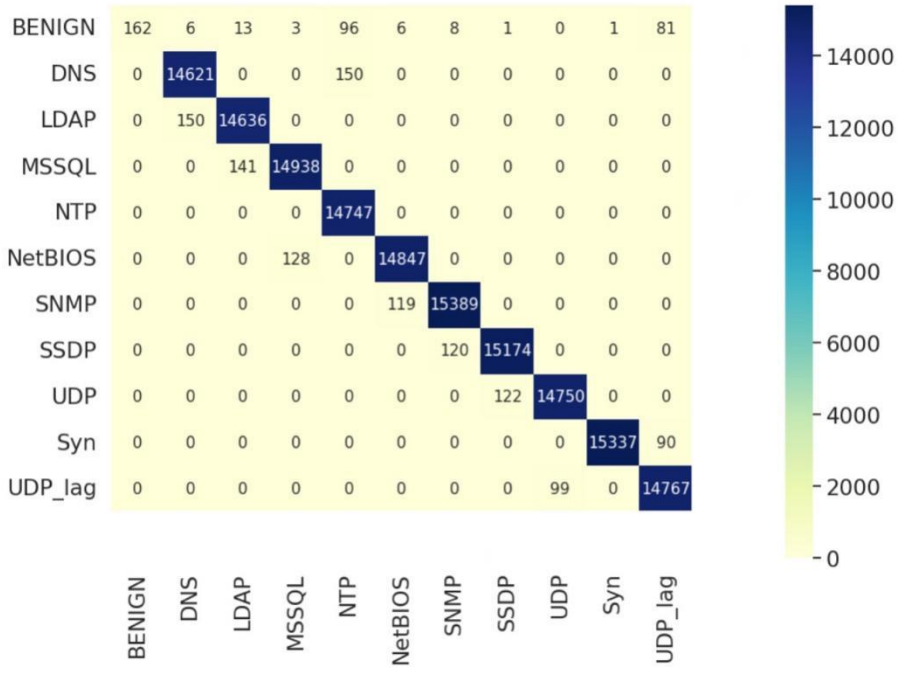

Figure 3. Confusion matrix of prediction results from decision tree.

4.3. Long short term memory

LSTM has the accuracy score, 98.199%. The table 3 and figure 4 show the F1 scores and confusion matrix of prediction results. Same as SVM and decision, this approach can also perfectly detect SYN. The F1-score of LSTM is 0.42 for BENIGN, which means LSTM cannot detect normal traffic sufficiently like decision tree. Other types of DDos traffic shows good F1 scores but NetBIOS, SNMP, UDP-lag and UDP have lower scores than the values in decision tree.

DOI: 10.54254/2755-2721/29/20230972

Table 3. Evaluation results of LSTM.

	Precision	Recall	F1 Score
Syn	1.00	1.00	1.00
NTP	0.99	1.00	0.99
MSSQL	0.99	0.99	0.99
LDAP	0.99	0.99	0.99
DNS	0.99	0.99	0.99
SSDP	0.99	0.99	0.99
NetBIOS	0.98	0.98	0.98
SNMP	0.98	0.98	0.98
UDP-lag	0.99	0.93	0.96
UDP	0.93	0.99	0.96
BENIGN	0.92	0.27	0.42
WebDDos	0.00	0.00	0.00

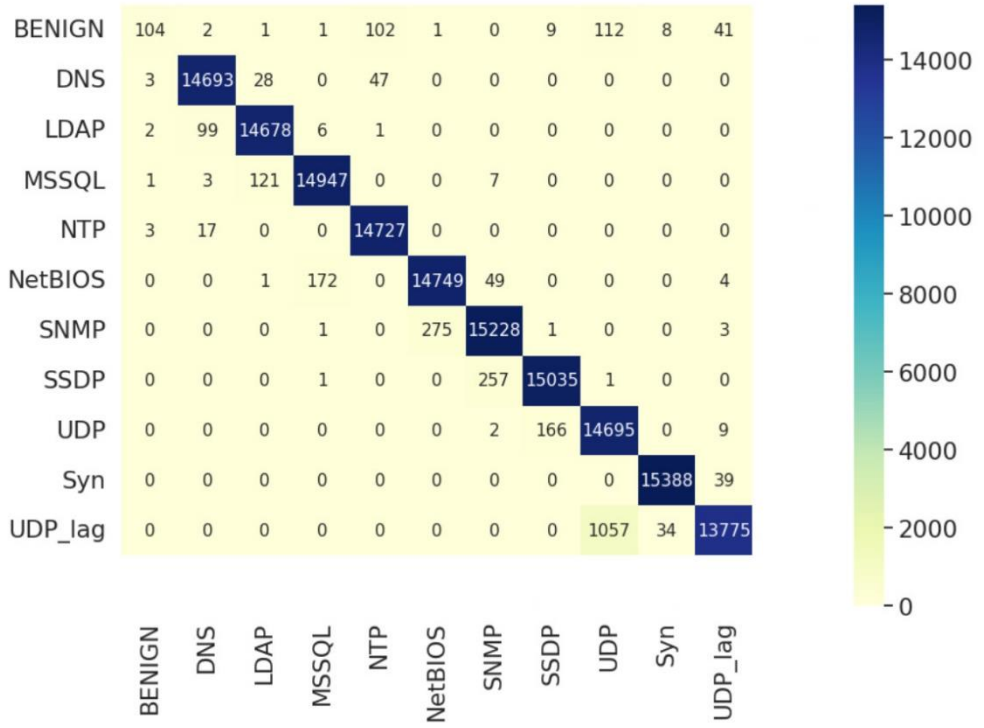

Figure 4. Confusion matrix of prediction results from LSTM.

4.4. Random forest

Random Forest shows the highest accuracy score which is 99.321%. The table 4 and figure 5 show the evaluation results of the prediction from Random Forest. It is clear that Random Forest has the best F1-score. This algorithm can perfectly detect SYN, NTP DDos attacks. Due to 0.96 score of BENIGN, Random Forest also has the best performance on the detection of normal and abnormal traffic. The rest types of DDos all have 0.99 F1-Score. It means that Random Forest can sufficiently detect these DDos attack traffic.

DOI: 10.54254/2755-2721/29/20230972

Table 4. Evaluation results of random forest.

	Precision	Recall	F1 Score
Syn	1.00	1.00	1.00
NTP	1.00	1.00	1.00
UDP-lag	0.99	0.99	0.99
SNMP	0.99	0.99	0.99
LDAP	0.99	0.99	0.99
NetBIOS	0.99	0.99	0.99
DNS	0.98	1.00	0.99
UDP	0.99	0.99	0.99
MSSQL	0.99	1.00	0.99
SSDP	0.99	0.99	0.99
BENIGN	0.92	1.00	0.96
WebDDos	0.90	0.68	0.78

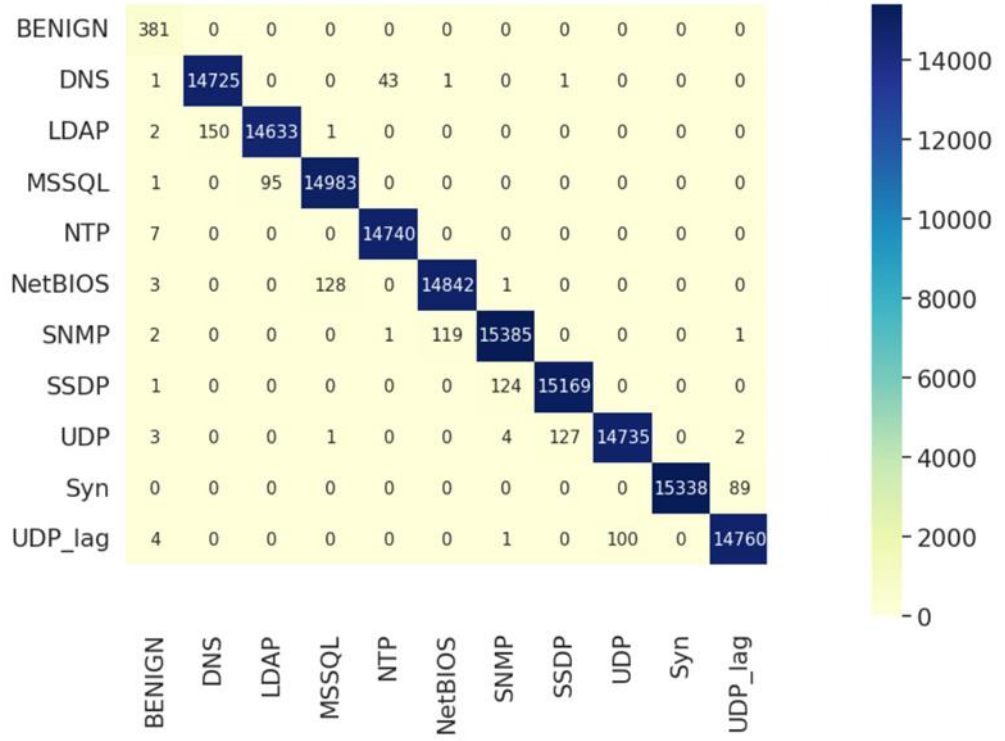

Figure 5. Confusion matrix of prediction results from Random Forest.

5. Conclusion

This study aimed to implement multiclass classification for detecting different types of DDoS attacks in IoT networks using SVM, Decision Tree, LSTM, and Random Forest models. The primary objective was to train these models to distinguish between various types of DDoS traffic and generate high F1-scores for all ten types. The results showed that all four models were able to detect the different types of DDoS traffic with relatively high accuracy, while SVM and Random Forest were particularly effective in distinguishing between normal and abnormal traffic. The Random Forest model had the best performance overall, with the highest accuracy score (99.321%) and F1-scores compared to the other models. Furthermore, this study suggests that this approach could be extended by incorporating more

DOI: 10.54254/2755-2721/29/20230972

types of DDoS traffic to develop a more comprehensive detection system for IoT networks. For example, the traffic data for DDoS attacks using TFTP could be added to the models. Moreover, while this study focused on detecting DDoS traffic in IoT networks, future research could explore developing defense mechanisms for different types of DDoS attacks and combine them with machine learning detection systems. In summary, this study demonstrates the effectiveness of machine learning models for detecting different types of DDoS attacks in IoT networks and highlights the potential for further research to expand the scope of these models and enhance their overall effectiveness.

References

[1] Iot-Analytics. (2022). Iot 2021 in review: The 10 most relevant iot developments of the year. https://iot-analytics.com/iot-2021-in-review/

[2] Wei, W., Yang, A. T., Shi, W., et al. (2016). Security in internet of things: Opportunities and challenges. In 2016 International Conference on Identification, Information and Knowledge in the Internet of Things (IIKI) (pp. 512-518). IEEE.

[3] Kolias, C., Kambourakis, G., Stavrou, A., et al. (2017). DDoS in the IoT: Mirai and Other Botnets. Computer, 50(7), 80-84.

[4] Yoachimik. (2023). Cloudflare DDoS Threat Report for 2022 Q4. Cloudflare. https://blog.cloudflare.com/ddos-threat-report-2022-q4/

[5] Microsoft Security. (2023). 2022 in review: DDOS attack trends and insights. https://www.microsoft.com/en-us/security/blog/2023/02/21/2022-in-review-ddos-attack-trends-and-insights/

[6] Sharafaldin, I., Lashkari, A. H., Hakak, S., et al. (2019). Developing Realistic Distributed Denial of Service (DDoS) Attack Dataset and Taxonomy. In 2019 International Carnahan Conference on Security Technology (ICCST) (pp. 1-8). IEEE.

[7] Kumari, P., & Jain, A. K. (2023). A comprehensive study of ddos attacks over IOT network and their countermeasures. Computers & Security, 127, 103096.

[8] Vishwakarma, R., & Jain, A. K. (2019). A survey of ddos attacking techniques and defence mechanisms in the IOT network. Telecommunication Systems, 73(1), 3-25.

[9] Suresh, M., & Anitha, R. (2011). Evaluating machine learning algorithms for detecting ddos attacks. In Advances in Network Security and Applications (pp. 441-452). Springer, Berlin, Heidelberg.

[10] Doshi, R., Apthorpe, N., & Feamster, N. (2018). Machine Learning DDoS Detection for Consumer Internet of Things Devices. In 2018 IEEE Security and Privacy Workshops (SPW) (pp. 29-35). IEEE.

[11] Tuan, T. A., Long, H. V., Son, L. H., et al. (2020). Performance evaluation of botnet ddos attack detection using machine learning. Evolutionary Intelligence, 13(2), 283-294.

[12] Chen, Y.-W., Sheu, J.-P., Kuo, Y.-C., et al. (2020). Design and Implementation of IoT DDoS Attacks Detection System based on Machine Learning. In 2020 European Conference on Networks and Communications (EuCNC) (pp. 122-127). IEEE.

[13] Gaur, V., & Kumar, R. (2021). Analysis of machine learning classifiers for early detection of ddos attacks on IOT devices. Arabian Journal for Science and Engineering, 47(2), 1353-1374.

[14] Raschka S, Mirjalili V. Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2[M]. Packt Publishing Ltd, 2019.

[15] Alves F R V, Vieira R P M. The Newton fractal's Leonardo sequence study with the Google Colab[J]. International Electronic Journal of Mathematics Education, 2019, 15(2): em0575.