Journal of Technology Research (JTR)
Volume 3, Special Issue, (2025), pp 595-601 ISSN 3005-639X
The 7th Conference of Engineering Sciences and Technology
(CEST-2025)

Supervised Machine Learning Approaches for Robust DDoS Detection in Cloud Environments

Sabria A. Bennaser[1], Haitham S. Ben Abdelmula[2], Abdusamea Omer[3], Ali Elghirani[4]

[1]Department of Computer Science, School of Basic Science, Libyan Academy for Postgraduate Studies, Misurata, Libya.

[2]Department of Computer Networks, College of Computer Technology Zawia, Zawia, Libya.

[3]Department of Computer Engineering, College of Engineering, Sabratha, Libya.

[4]Faculty of Information Technology, Libyan International Medical University, Benghazi, Libya.

*Corresponding author email: hsaa8383@gmail.com

Received: 14-10-2025 | Accepted: 02-12-2025 | Available online: 25-12-2025 | DOI:10.26629/jtr.2025.55

ABSTRACT

In today's landscape, the widespread adoption of cloud computing has been accompanied by a corresponding increase in security vulnerabilities, with Distributed Denial-of-Service (DDoS) attacks posing one of the most serious challenges by overwhelming resources such as CPU power, memory, and network bandwidth, thereby disrupting services for legitimate users. Detecting DDoS attacks in cloud environments is particularly difficult due to the similarity between malicious and legitimate traffic, often originating from numerous geographically dispersed sources. This study evaluates the effectiveness of five supervised machine learning algorithms Random Forest (RF), Decision Tree (DT), Support Vector Machine (SVM), k-Nearest Neighbours (KNN), and Naïve Bayes (NB) for detecting DDoS attacks in cloud computing environments using the publicly available Software Defined Networking (SDN) DDoS Attack Dataset. Comprehensive preprocessing including normalization, feature selection, and Synthetic Minority Oversampling Technique (SMOTE) was applied, along with rigorous regularization strategies to mitigate overfitting. Experimental results demonstrate that Random Forest achieved the highest balanced performance (95% accuracy, 96% precision, 95% recall), followed by KNN (94%), SVM (93%), DT (92%), and Naïve Bayes (91%). These findings confirm the potential of machine learning for reliable DDoS detection while emphasizing the importance of proper model regularization to ensure generalizability. Future work should explore larger datasets, real-time traffic analysis, and hybrid models to further enhance robustness.

Keywords: Machine Learning, Supervised Algorithms, DDoS Attacks, Cloud Computing, SDN.

أساليب التعلم الآلي تحت الإشراف للكشف القوي عن هجمات حجب الخدمة (DDoS) في بيئات الحوسبة السحابية

صبرية أحمد بن نصر[1]، هيثم صالح بن عبد المولى[2]، عبدالسميع عمر[3]، علي الغرياني[4]

[1] قسم علوم الحاسوب، كلية العلوم الأساسية، الأكاديمية الليبية للدراسات العليا، مصراتة، ليبيا

[2] قسم شبكات الحاسوب، كلية تقنية الحاسوب الزاوية، الزاوية، ليبيا

[3] قسم هندسة الحاسوب، كلية الهندسة، صبراتة، ليبيا

[4] كلية تقنية المعلومات، الجامعة الطبية الليبية الدولية، بنغازي، ليبيا

ملخـــــــص البحـــــــث

في المشهد الحالي، صاحب الانتشار الواسع للحوسبة السحابية زيادة مقابلة في الثغرات الأمنية، حيث تشكل هجمات الحرمان الموزع من الخدمة (**DDoS**) أحد أخطر التحديات من خلال استنزاف موارد مثل قدرة المعالج والذاكرة وعرض النطاق الترددي للشبكة، مما يعطل الخدمات للمستخدمين الشرعيين. يعد كشف هجمات **DDoS** في البيئات السحابية صعبًا بشكل خاص بسبب التشابه بين الحركة الضارة والحركة المسموحة ، التي غالبًا ما تنبع من مصادر متعددة موزعة جغرافيًا. تقيم هذه الدراسة فعالية خمسة خوارزميات تعلم آلي خاضع للإشراف **Random Forest (RF), Decision Tree (DT), Support Vector Machine (SVM), k−Nearest Neighbours (KNN), and Naïve Bayes (NB)** لاكتشاف هجمات **DDoS** في بيئات الحوسبة السحابية باستخدام مجموعة بيانات هجوم **SDN DDoS** المتاحة للعموم. تم تطبيق معالجة مسبقة شاملة شملت التطبيع، واختيار الميزات، وتقنية إعادة العينة الاصطناعية للأقلية (**SMOTE**)، إلى جانب استراتيجيات تنظيم صارمة للتخفيف من **overfitting**. تظهر النتائج التجريبية أن خوارزمية **Random Forest** حققت أعلى أداء متوازن (دقة 95%، دقة تنبؤ 96%، استدعاء 95%)، تليها **KNN** (94%)، ثم **SVM** (93%)، فـ**DT** (92%)، وأخيرًا **Naïve Bayes** (91%). تؤكد هذه النتائج إمكانات التعلم الآلي في الكشف الموثوق عن هجمات **DDoS** مع التأكيد على أهمية التنظيم السليم للنماذج لضمان القدرة على التعميم. وكتوصيات لاعمال مستقبلية يمكن العمل بمجموعات بيانات أكبر، وتحليل حركة المرور في الوقت الفعلي، والنماذج الهجينة لتعزيز المتانة أكثر.

الكلمات الدالة: التعلم الآلي، الخوارزميات تحت الإشراف، هجمات حجب الخدمة، الحوسبة السحابية، الشبكات المعرفة بالبرمجيات (SDN).

1. INTRODUCTION

Cloud computing has emerged as a transformative technology, offering a flexible and on-demand pool of configurable computing resources that can be accessed with minimal management effort or service provider intervention [1]. Despite its numerous advantages such as scalability, cost efficiency, and accessibility cloud platforms face significant security challenges. Unlike traditional infrastructures that rely on static and well-defined perimeters, cloud environments operate on shared infrastructures, support multi-tenancy, and enable dynamic resource allocation, all of which introduce unique vulnerabilities and increase the potential attack surface [2]. Among the diverse range of security threats, Distributed Denial-of-Service (DDoS) attacks remain among the most severe and disruptive. In these attacks, malicious actors flood target servers with excessive traffic, often leveraging networks of compromised devices (botnets) to exhaust system resources such as CPU, memory, and bandwidth. Since attackers frequently mimic legitimate traffic patterns and utilize geographically dispersed sources, differentiating between normal and malicious requests becomes a complex task [2].

Moreover, conventional network-layer defense mechanisms are often ineffective against modern application-layer DDoS attacks, highlighting the urgent need for more adaptive, intelligent, and data-driven detection strategies capable of operating efficiently in cloud environments.

While deep learning approaches have gained recent attention in cybersecurity, traditional supervised learning methods remain highly relevant for DDoS detection due to their interpretability, computational efficiency, and proven effectiveness on structured network traffic data.

To address these challenges, this study makes three primary contributions. First, it presents a comparative evaluation of five supervised machine learning algorithms Random Forest (RF), Decision Tree (DT), Support Vector

Machine (SVM), k-Nearest Neighbors (KNN), and Naïve Bayes (NB) using a publicly available SDN-based DDoS dataset. Second, it implements comprehensive preprocessing techniques, including normalization, feature selection, and the Synthetic Minority Oversampling Technique (SMOTE), to mitigate class imbalance and improve overall model performance. Third, it provides empirical insights into the relative strengths and weaknesses of each algorithm, offering practical guidance for developing effective and scalable DDoS detection systems tailored for cloud computing environments.

2. DDoS ATTACK PROCESS

DDoS Attack aims to disrupt the services of a specific network or server by flooding it with multiple requests from multiple sources at the same time as shown in Figure 1. The attacker hacks a large number of devices (such as computers, smartphones, etc.) and turns them into bots. These hacked devices are part of a network called a botnet.

Attack Routing: The attacker issues commands to these bots to send requests or data to the target server at the same time.

Server Flooding: These concurrent requests flood the server, making it unable to handle legitimate requests from actual users. This results in slow performance or complete cessation of service.

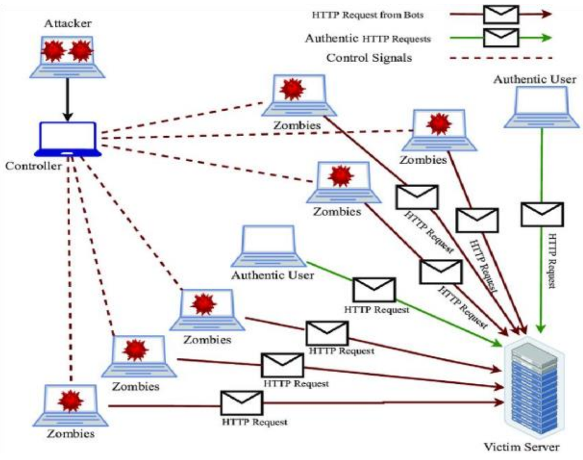

Fig 1. DDoS Attack Process.

3. RELATED WORKS

Numerous research studies have been done on security enhancement and have proposed several techniques to attain the desired security level. Tan, Liang et al. [3] introduced an innovative security framework tailored for mitigating DDoS attacks within SDN environments. Their model comprises two key modules leveraging machine learning (ML) algorithms. The data-processing module utilizes the K-Means algorithm to optimize feature selection, while the detection module employs the k-nearest neighbor (KNN) algorithm to identify attack flows. In comparison to approaches like distributed Self-Organizing Map (SOM) and entropy-based methods, their model achieves an impressive accuracy of 98.85% with a recall rate of 98.47%. Shaw et al. [4] they introduced an evolutionary approach for classifying DDoS attack traffic within an SDN framework. This model employs an SVM algorithm for the classification of malicious traffic, utilizing genetic algorithms (GA) to optimize SVM parameters. Kernel Principal Component Analysis (KPCA) is selected as a method to enhance feature selection. The model's classification performance is evaluated using datasets containing UDP flood, HTTP streaming, and regular traffic to assess its accuracy. Experimental results demonstrate that the combined approach achieves an accuracy of 98.9%. Mishra et al. [5] they proposed a system based on supervised machine learning To detect and prevent DDoS attacks on the server via the cloud and extract statistical features. Using Naïve Bayes (NB), Nearest Neighbours (KNN) and Random Forest (RF) classifiers, the results showed that the proposed approach can detect DDoS attacks with almost high accuracy (99.68%) using RF and low Fake positives. Yassin et al. [6], they proposed a method based on Naïve Bayes (NB) and K-means clustering to detect DDoS attacks. The K-means clustering method groups traffic data exhibiting similar behaviours, while the Naïve Bayes algorithm classifies the clustered data into normal and attack traffic categories. Ajeetha et al. [7] they proposed a method to detect spreader denial-of-service attacks by analyzing traces in traffic flows. A confusion matrix was constructed based on these traces, and two classifiers, Naive Bayes and Random Forest, were employed to classify traffic as either normal or abnormal, using profiles derived from existing datasets of normal and attack behaviours. Naive Bayes algorithm outperformed the Random Forest algorithm in terms of accuracy and effectiveness. M NALAYINI et al. [8] this study explores machine learning algorithms for detecting distributed denial-of-service (DDoS) attacks, utilizing the NSL-KDD dataset. Two techniques are employed: Learning Vector Quantization (LVQ) for classification and Principal Component Analysis (PCA) for dimensionality reduction. Each approach selects specific characteristics to detect DDoS attacks effectively. Decision Tree (DT), Naïve Bayes (NB), and Support Vector Machine (SVM) classifiers are applied and compared in terms of classification performance. Among these, the LVQ-based Decision Tree stands out for its superior ability to identify attacks compared to other DT variants.

4. METHODOLGIES

4.1 Description of Dataset

This study uses the public SDN DDoS Attack Dataset [9], which was created in a Software Defined Networking (SDN) environment and made available for research purposes. The dataset contains 104,345 traffic flow records and includes 23 features representing various network characteristics. Each record is labelled as either normal traffic or attack traffic, allowing for supervised learning. The dataset covers different network protocols, including TCP, UDP, and ICMP, providing a balanced view of real-world network behaviour under both normal and attack conditions.

4.2 Data Preprocessing

Data preprocessing is a critical step to ensure the accuracy and reliability of machine learning models. The process begins with an examination of the dataset to identify missing values and inconsistencies. Features that do not contribute meaningfully to classification are removed after analyzing their correlation with the output labels. Since the dataset contains an uneven distribution of attack and normal traffic, the Synthetic Minority Oversampling Technique (SMOTE) [12] is applied to balance the data. This technique generates synthetic examples of the minority class to reduce bias during training. All numerical features are then normalized to ensure uniform scaling and to improve the learning performance of the classifiers. Finally, the processed dataset is divided into training and testing subsets using a 70:30 ratio, allowing for an objective evaluation of model performance.

4.3 Machine Learning Algorithms

Machine learning techniques are used to analyze system performance and detect abnormal network behaviours. In this study, five supervised learning algorithms are implemented: Random Forest, Decision Tree, Support Vector Machine, k-Nearest Neighbours, and Naïve Bayes. Each method has distinct characteristics and advantages, as summarized below.

Random Forest (RF)

Random Forest is an ensemble learning method that combines multiple decision trees to produce more accurate and stable predictions [10]. Each tree is trained on a random subset of the data and predictor variables, reducing overfitting and improving generalization.

Decision Tree (DT)

A Decision Tree consists of internal nodes that represent decision rules and leaf nodes that represent output classes. It supports both categorical and numerical data, making it suitable for classification problems in network analysis [10].

Support Vector Machine (SVM)

SVM is a powerful supervised learning algorithm used for classification and regression. It identifies the optimal hyperplane that separates data points into two distinct classes (e.g., normal vs. attack) while maximizing the margin between them [10].

k-Nearest Neighbours (KNN)

KNN is a simple yet effective algorithm that classifies data based on the majority label of its k nearest neighbours. It is widely used in anomaly detection due to its ability to adapt to local data structures [10].

Naïve Bayes (NB)

Naïve Bayes is a probabilistic classifier that applies Bayes theorem under the assumption that features are independent. It is computationally efficient and performs well even with limited data, making it useful for rapid classification tasks [10].

4.4 Evaluation Metrics

To assess model performance, several standard evaluation metrics are used, including accuracy, precision, recall, and F1-score. These metrics are derived from the confusion matrix, which compares the predicted and actual class labels [11]. Accuracy measures the proportion of correct predictions, while precision focuses on how many of the positive predictions were correct. Recall evaluates how effectively the model identifies true attack instances, and the F1-score provides a balanced measure that combines precision and recall. All models are trained and tested under identical conditions to ensure a fair comparison of their detection capabilities.

5. APPLICATION AND RESULTS

In this study, five machine learning algorithms including decision tree, naive bayes, random forest, support vector machines, and k nearest neighbors are applied to classify and detect abnormal behavior in network traffic. Finally, different performance evaluation metrics called accuracy, precision, recall, and F1-score were used to assist the performance of those algorithms by using the confusion matrix.

5.1 Confusion Matrix

It is the easiest way to determine the performance of a classification model by comparing how many positive instances were correctly/incorrectly classified and how many negative instances were correctly/incorrectly classified. In a Confusion Matrix, the rows represent the actual labels and the columns represent the predicted labels [11].

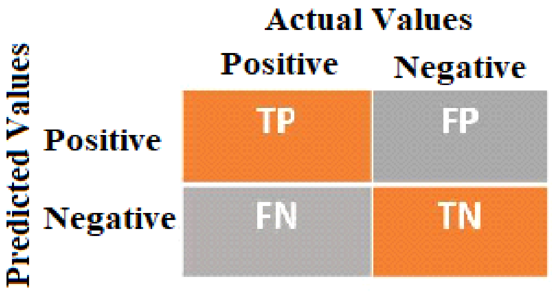

Fig 2. Confusion Matrix.

5.2 Performance Metrics

Performance Metrics Accuracy. Accuracy represents the proportion of correctly classified instances relative to the total number of predictions made by the model. While it is one of the most commonly reported metrics, accuracy alone can be misleading particularly in the presence of class imbalance, where one class significantly outnumbers the other. In such cases, a model may appear to perform well simply by favoring the majority class. Accuracy is computed as:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively. Precision. Precision measures the proportion of correctly predicted positive instances out of all instances predicted as positive. In other words, it quantifies how reliable the positive predictions are. High precision indicates a low rate of false positives, which is crucial in contexts where false alarms carry significant cost. It is defined as:

$$\boldsymbol{Precision} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FP}}$$

Recall. Recall, also known as sensitivity or true positive rate, quantifies the proportion of actual positive instances that the model successfully identifies. It reflects the model s ability to capture all relevant positive samples and is particularly important when missing a positive instance is costly. It is calculated as:

$$\boldsymbol{Recall} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FN}}$$

F1 Score.

The F1 Score provides a balanced measure between precision and recall by taking their harmonic mean. It is especially useful in imbalanced datasets, where a single metric (such as accuracy) may not fully capture model performance. The F1 Score is defined as:

$$\boldsymbol{F1} = \boldsymbol{2} \cdot \frac{\boldsymbol{precision \cdot recall}}{\boldsymbol{precision+recall}}$$

Together, these metrics offer a comprehensive view of classification performance, balancing the trade-offs between precision, recall, and overall predictive accuracy.

5.*3 Performance Comparison*

The performance of each model is examined by applying each algorithm to the dataset, applying the four metrics, obtaining the results and comparing them with the results obtained in the paper under study. The comparison results appear in the table below.

Table 1 : Performance Analysis of different approaches**.**

	Accuracy	Precision	Recall	F1 Score
NB	0.91	0.92	0.90	0.91
KNN	0.94	0.95	0.94	0.94
RF	0.95	0.96	0.95	0.95
DT	0.92	0.93	0.92	0.92
SVM	0.93	0.94	0.93	0.93

6. Conclusions

Detecting Distributed Denial of Service (DDoS) attacks is a critical issue, as these attacks can significantly disrupt cloud services. Machine learning models have the ability to effectively identify such attacks. This study focuses on accurately detecting distributed DDoS attacks, using data specifically related to these incidents. Various machine learning methods have been used, including decision tree (DT), support vector machine (SVM), random forest (RF), k-nearest neighbours (KNN), and Naïve Bayes (NB), along with feature selection algorithms for classification. Conducting a balancing process for the data sets, all methods showed high accuracy in classifying DDoS attacks, higher than the percentages they obtained in the comparative search, with Random Forest achieving the highest accuracy rate of 95%, followed by k-Nearest Neighbours (94%), Support Vector Machine (93%), Decision Tree (92%), and Naïve Bayes (91%).

.

REFERENCES

[1] Kadhim, M.A.; Radhi, A.M. Machine Learning Prediction of Brain Stroke at an Early Stage. Iraqi Journal of Science, 2023.

[2] Wani, A.R.; Rana, Q.P.; Pandey, N. Machine Learning Solutions for Analysis and

Detection of DDoS Attacks in Cloud Computing Environment. International Journal of Engineering and Advanced Technology, 2020, 9(3), 2205 2209.

[3] Tan, L.; Pan, Y.; Wu, J.; Zhou, J.; Jiang, H.; Deng, Y. A New Framework for DDoS Attack Detection and Defense in SDN Environment. IEEE Access, 2020.

[4] Sahoo, K.S.; Tripathy, B.K.; Naik, K.; Ramasubbareddy, S.; Balusamy, B.; Khari, M.; Burgos, D. An Evolutionary SVM Model for DDoS Attack Detection in Software Defined Networks. IEEE Access, 2020.

[5] Mishra, A.; Gupta, B.B.; Peraković, D.; Peñalvo, F.J.G.; Hsu, C.H. Classification-Based Machine Learning for Detection of DDoS Attack in Cloud Computing. In Proceedings of the IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, January 2021; pp. 1 4.

[6] Yassin, W.; Udzir, N.I.; Muda, Z.; Sulaiman, M.N. Anomaly-Based Intrusion Detection through K-Means Clustering and Naïve Bayes Classification. 2013.

[7] Ajeetha, G.; Priya, G.M. Machine Learning Based DDoS Attack Detection. In Proceedings of Innovations in Power and Advanced Computing Technologies (i-PACT), 2019, 1, 1 5.

[8] Nalayini, M.C.; Katiravan, J. Detection of DDoS Attack Using Machine Learning Algorithms. Journal of Emerging Technologies and Innovative Research, 2022, 9(7).

[9] DDoS SDN Dataset. Available online: https://www.kaggle.com/datasets/aikenkazin/ddos-sdn-dataset.

[10] Tonkal, Ö.; Polat, H.; Başaran, E.; Cömert, Z.; Kocaoğlu, R. Machine Learning Approach Equipped with Neighborhood Component Analysis for DDoS Attack Detection in Software-Defined Networking. Electronics, 2022, 10 (11)

[11] Hastie, T.; Tibshirani, R.; Friedman, J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed.; Springer: New York, 2009.

[12] Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research 2002, 16, 321-357.