









Supervised learning for Attack Detection in Cloud

Animesh Kumar, Sandip Dutta and Prashant Pranav*



Department of Computer Science and Engineering, Birla Institute of Technology, Mesra, Ranchi, Jharkhand, India

E-mail/Orcid Id:

AKS,  animeshkumarce@gmail.com,  <https://orcid.org/0009-0007-3679-8025>; SD,  sandipdutta@bitmesra.ac.in,  <https://orcid.org/0000-0002-3932-3048>; PP,  prashantpranav19@gmail.com,  <https://orcid.org/0000-0002-3932-3048>

Article History:

Received: 29th Mar., 2023

Accepted: 22nd Jun., 2023

Published: 30th Jul., 2023

Keywords:

Cloud Attack, Cloud Computing, Machine Learning, Supervised Learning, Security Issues, Support Vector Machine

Abstract: In this study, we approach a supervised learning algorithm to detect attacks in cloud computing. We categorize "Normal" and "Attack" statuses on the dataset. The model evaluation process uses the kappa statistic, the F1-score, recall, accuracy, and precision. The system has a very high detection and efficiency rate, with a detection rate of over 99%. A total of 9594 cases and 44 distinct columns are included in the dataset. The study's results were displayed using a ROC curve and a confusion matrix. This study focuses on implementing a supervised learning algorithm for detecting attacks in cloud computing environments. The main objective is distinguishing between "Normal" and "Attack" statuses based on a carefully curated dataset. Several metrics, such as the kappa statistic, F1-score, recall, accuracy, and precision, are employed to evaluate the model's performance. The dataset utilized in this research comprises 9594 cases and encompasses 44 distinct columns, each representing specific features relevant to cloud computing security. Through a rigorous evaluation process, the algorithm demonstrates exceptional efficiency, achieving a remarkable detection rate of over 99%. Such high accuracy in identifying attacks is crucial for ensuring the integrity and security of cloud-based systems. The significance of this study lies in its successful application of a supervised learning approach to tackle cloud computing security challenges effectively. The model's high detection rate and efficiency indicate its potential for real-world deployment in cloud-based systems, contributing to enhanced threat detection and mitigation. These results hold promising implications for bolstering the security measures of cloud computing platforms and safeguarding sensitive data and services from potential attacks.

Introduction

Traditional Cloud computing is the on-demand provisioning of computer resources over the Internet, including servers, storage, databases, networking, software, etc., over the Internet (Butt et al., 2023). It can be easily scaled up or down based on demand, allowing for flexibility and cost optimization. The responsibility for infrastructure maintenance, security patches, and updates lies with the cloud provider. Amazon Web Service, Microsoft Azure, Google Cloud Platform, IBM Cloud, and Oracle Cloud are popular cloud service providers. There are three types of cloud deployment models. The public cloud is available for public use. Private Cloud, where Infrastructure and services are

dedicated to a single organization and can be hosted internally or externally. Hybrid Cloud combines public and private clouds, allowing data and applications to be shared between them (Jain and Rajak, 2023).

Figure 1 explains the collaborative efforts of cloudlets, brokers, data centers, and services to provide efficient and accessible cloud services. The process begins when a user initiates a request for a specific cloud service or application. The request is forwarded if a cloudlet is nearby, enabling faster processing and reduced latency. Based on the evaluation, the broker selects the most suitable cloud service provider and communicates the user's request and service selection to the chosen data center. The data center receives the request, allocates the

*Corresponding Author: prashantpranav19@gmail.com



necessary computing resources, and executes the requested service or application. This collaborative working of cloudlets, brokers, data centers, and services optimizes the delivery of cloud services, reduces latency, enhances user experience, and provides users with efficient access to a wide range of computing resources and applications. Some of the security attacks are as follows.

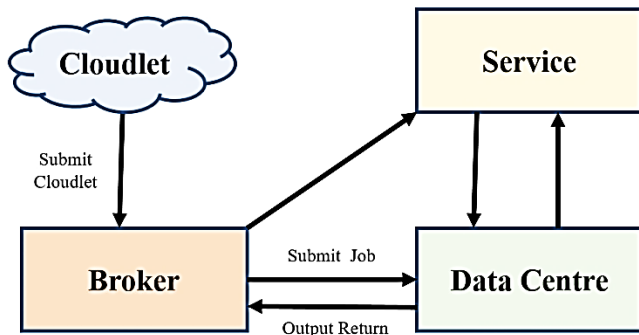


Figure 1. Cloud Computing Architecture

Denial of Service (DoS) Attacks

These attacks can disrupt the normal functioning of cloud services, leading to downtime, loss of access, and potential financial and reputational damages for both the cloud service provider and its customers (George et al., 2023; Gemmer et al., 2023; Clemens et al., 2023; Ashlam et al., 2023; Anita et al., 2023). Side Channel Attacks: These attacks take advantage of the physical characteristics, timing, power consumption, etc., of systems to extract sensitive information (Yan et al., 2015; Sahi et al., 2017; Agarwal et al., 2019; Gopinath et al., 2023). Man-In-The-Middle Cryptographic Attacks, the attacker captures or alters the transmitted data, including sensitive user credentials, financial information, or critical application data (Lu et al., 2021; Ha et al., 2022; Zhang et al., 2019; Sultan et al., 2022). Strong Authentication and Access Controls, Data Encryption, Regular Security Updates, and Patching Network Security (Ma et al., 2023; Gong et al., 2020) measures like Implementing firewalls, intrusion detection & prevention systems (Radhakishan et al., 2011; Ren et al., 2020a,b). Moreover, network segmentation protects cloud networks from unauthorized access and malicious activities.

Machine Learning (ML) is a subset of artificial intelligence (Joshi et al., 2023). that involves the development of algorithms and models that enable computers to learn from data and make predictions. Algorithms that can automatically learn and improve from experience instead of relying on instructions. Image and speech recognition, natural language processing (NLP) (Dash et al., 2023; Khurana et al., 2023), anomaly

detection, predictive analytics, etc., done using ML (Kreuzberger et al., 2023; Kwekha-Rashid et al., 2023). Supervised learning is a machine learning approach that involves training models on labeled data, where each data point is associated with a known outcome. Supervised learning aims to learn a mapping function that can accurately predict or classify new, unseen instances based on their features (Yu et al., 2023; Wu et al., 2023). Regression and Classification are two types of supervised learning. Classifiers like simple vector machine (SVM) (Kurani et al., 2023), Random Forest (Bicego et al., 2023), Decision Tree (Utukuru et al., 2023), Logistic regression (Wang et al., 2023), Naïve Bayes (Saleh et al., 2023), Xtreme Gradient Boosting (XGBoost) (Iban et al., 2023), K-Nearest Neighbour (K-NN) (Mohy-Eddie et al., 2023) etc. are an example of supervised learning.

(Aldhyani et al., 2022) suggested EDoS attacks in cloud computing using ML and DL methods using SVM, KNN, RF, and Deep learning methods.(Arunkumar et al., 2023) Proposed that Gannet Optimization Algorithm-based hybrid SVM-ELM mitigates attacks in the cloud. Matlab software is used for simulation using CICIDS2017 datasets and Proposed a DL fusion-based method to solve DDoS issues in cloud computing. More optimized DL methods can be deployed for better cloud detection. (Patel et al., 2022) reviewed the DL method for attack detection in the network. (Verma et al., 2023) proposed the ReputE method for DDoS attack detection in IoT and fog computing. Classifiers will be designed to counter future live IoT network traffic attacks; proposed a Text-Mining approach for accident analysis in steel plants to reduce human involvement in accident-prone areas. Text mining is done in two phases. Four years of data are used for model building and its analysis. However, there is scope for future enhancements using different ML classifiers. The author uses SVM, ANN, NB, K-NN, and RF Classifiers for predicting and analyzing injury severity (Sarkar et al., 2020). In future studies, accidents are also predicted using the time series method. Some of the other relevant papers (Sarkar et al., 2019; Pramanik et al., 2021; Das et al., 2022; Paramanik et al., 2022; Bag et al., 2023; Dey et al., 2023).

This research aims to develop and implement supervised machine learning algorithms to detect intrusion attempts in a cloud network. The objective is to accurately classify instances as either 'Normal' or 'Attack' to enhance the security of cloud computing. This work aims to design and implement an anomaly detection-based network intrusion detection system for a cloud computing network that can detect as many potential security issues as feasible.

Table 1. Pseudocode for Logistic Regression

```

function Logistic_Regression_Train(dataset)
weights = initialize_weights ()
bias = initialize bias ()
for iteration = 1 to num_iterations do:
gradients = compute_gradients (dataset, labels, weights, bias)
weights -= learning_rate * gradients
bias -= learning_rate * compute_bias_gradient(gradients)
return weights, bias
function Logistic_Regression_Predict (dataset, weights, bias):
predictions = []
For instance, In the Dataset
prediction = sigmoid (dot product (weights, instance) + bias)
predictions. Append(prediction)
return predictions
function sigmoid(x):
return 1 / (1 + exp(-x))

```

Research Questions

- Q1. What precautions must a user consider before going for cloud computing?
- Q2. How to secure the data while transferring to the Cloud?
- Q3. How to make sure data stored in the Cloud is secured?

Contribution

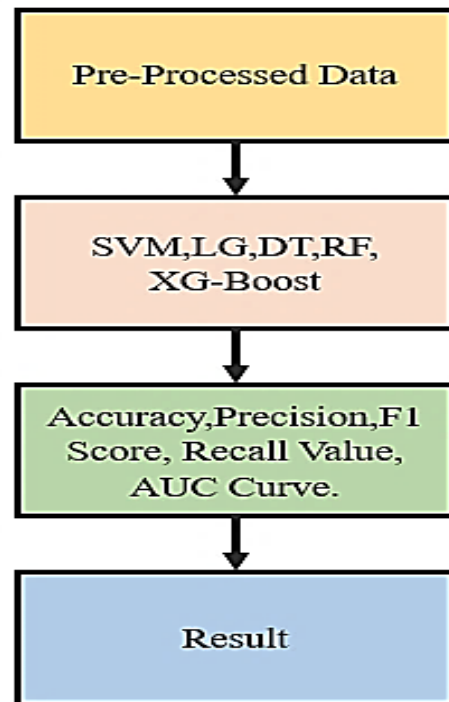
This study focuses on the hidden security attacks of CSPs that affect the quality of services resulting in much wastage of cloud resources and client money as the cloud work on a “Pay per basics model.” Detecting cloud attacks is very difficult as it involves a large set of traffic in real-time. Our smart model can help separate both normal and abnormal packets (attacks) using different classifiers of ML from the network, which is this paper's main contribution. In this experiment, we use an actual dataset from the cloud server. The result of this testing is equated with different standing systems to prove this system's durability and efficiency.

Materials and Method

Table 1 includes initializing the weights and bias, iterating over the number of iterations, computing gradients, updating the weights and bias using the learning rate, and using the sigmoid function for prediction. The `dot product` function calculates the dot product of the weights and an instance of the dataset.

Figure 2 shows the working of the proposed model. The first dataset is processed. The dataset is simulated on MATLAB 2023(a) using supervised Machine learning classifiers like LG, SVM, DT, RF, XG-Boost, etc. Parameters such as accuracy, precision, F1 score, and

Recall values were derived. The system with 16 GB RAM, 1 TB ROM, and

**Figure 2. Proposed methodology flowchart**

MATLAB version R2023 (a) is used to perform these experiments. A private cloud was used as the setting for the creation of the dataset. The private cloud infrastructure was set up with the help of a KVM type-1 hypervisor and an Open Nebula (5.12 version) cloud management platform. On cloud-based virtual machines, a script was run to generate a synthetic workload replicating the actual cloud model in real-time. We split our dataset in the ratio of 70: 30 during data pre-processing. The dataset is then prepared for training and testing by removing duplicates and outliers. We removed some extra features from the dataset to reduce the time

Table 2. Feature selection for attack status

LAST_POLL	rxbytes_slope	rxpackets_slope	txpackets_slope	timesys_slope	Status
1604624102	87.8402	27.2996	89.9974	17.8787	Attack
1604624071	87.9098	28.0725	89.9974	18.4349	Attack
1604624041	88.3794	30.3791	89.9974	34.5923	Attack
1604624012	88.0519	29.5388	89.9974	18.4349	Attack
1604623982	87.9098	28.0725	89.9974	18.4349	Attack
1604623952	87.9098	28.0725	89.9969	18.4349	Attack
1604623922	88.0114	29.5388	89.9973	33.6901	Attack
1604623892	88.0519	29.5388	89.9974	18.4349	Attack
1604623862	88.9491	37.7468	89.9971	17.8787	Attack
1604623831	87.9098	28.0725	89.9959	18.4349	Attack
1604623772	87.9546	28.0725	89.9974	18.4349	Attack
1604623742	87.9098	28.0725	89.9970	33.6901	Attack
1604623712	88.0114	29.5388	89.9968	18.4349	Attack

required to process the data. The irrelevant features and unused variables, such as the Time Stamp, Virtual Machine Identity, Unique Domain Identifier, and Domain Name, were removed. A distinct dataset, including characteristics, was developed following pre-processing. The subsequent tests were done on this dataset. In total, the dataset contains 9594 cases and 44 different columns. The first four columns of the table are used to hold metadata, which includes the epoch time, the virtual machine ID, the domain name, and the domain identifier. Two columns provide specific information about the available network, RAM, and disk space. In this table's last column, record whether the target is currently being attacked or functioning normally. Datasets are trained using a classifier like SVM, RF, LR, DT, K-NN, XG-Boost, and NB. Accuracy, Precision, Recall, F1-score, and Kappa statistics are all evaluated for each model.

Result and Discussion

Table 2 and Table 3 show the selected features, such as the last poll, rxbytes_slope, and txpackets_slope, with their values representing Attack and Normal status, respectively.

LAST_POLL

This column represents the date of the last poll, and rxbytes_slope represents the slope or rate of change of received bytes over time. rxpackets_slope represents the rate of change of received packets over time.

txpackets_slope: Column represents the rate of change of transmitted packets over time.

timesys_slope: This column represents the rate of change of system time over time. Status represents the device

being monitored. The values in each column would reflect the respective characteristics at that time.

Table 3 shows the categories of Normal from the dataset. In this research, the accuracy of the four different ensemble classifiers is analyzed and evaluated using the area under the curve as the metric of choice performance under comparison between imbalanced data and oversampled data.

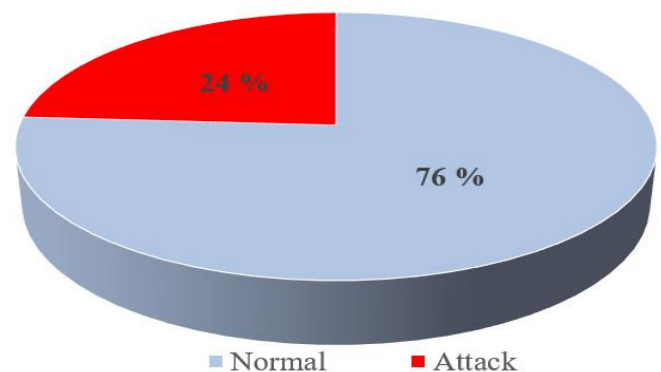


Figure 3. Distributed Targeted Variables

Fig 3 explains data Categorization into normal or under attack in pie chart form. Attack equals 2306 while Normal equals 7288.

Figure 4 shows a strong correlation between the status and the pace at which packets are sent over the network. It illustrates the link between features. The present state is significantly related to the amount of data transmitted over the network. The status positively correlates with the following metrics: the number of network packets received per second, the number of network bytes received per second, the number of network bytes transferred per second, the usage of kernel space, and user space.

Table 3. Feature selection for Normal status

LAST_POLL	rxbytes_slope	rxpackets_slope	txpackets_slope	timesys_slope	Status
1604455173	88.20650	30.14140	24.3045	89.9850	Normal
1604455142	87.87080	27.34990	15.9061	89.8986	Normal
1604455113	87.88650	27.29960	32.8285	89.9897	Normal
1604455082	87.87600	27.40760	14.2360	89.8741	Normal
1604455055	87.72410	25.82100	22.7510	89.9864	Normal
1604455024	87.71280	25.71000	18.4349	89.9685	Normal
1604454997	88.10170	29.74490	23.1986	89.9864	Normal
1604454962	87.87600	27.40760	12.5288	89.9580	Normal
1604454935	87.72860	25.86640	16.8584	89.9829	Normal
1604454902	87.79740	26.56510	21.8014	89.9818	Normal
1604454580	87.79740	26.56510	32.2756	89.9887	Normal
1604454542	87.87080	27.34990	17.8533	89.9324	Normal
1604454513	87.79740	26.56510	33.6901	89.9907	Normal

Figure 5 explains the statistical technique of classifying objects, data points, or clusters based on their similarities or dissimilarities. Cluster characteristics and differences between clusters can be analyzed to achieved an accuracy of 99.04%, better than many other proposed models. Our model compares with (Aldhyani et al., 2022; Fazlullah et al., 2023; Sagarkumar, 2023; GSR et al., 2023), which have an accuracy of 86.23%,96.53%,

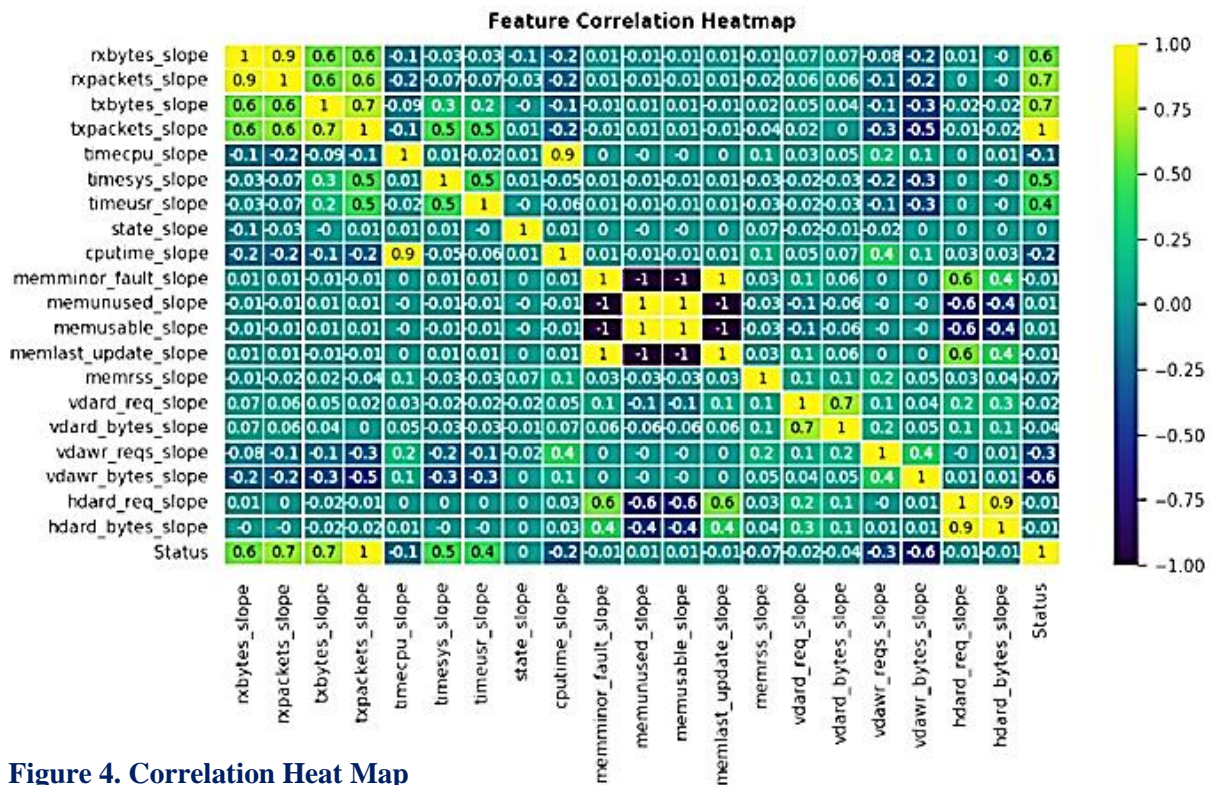


Figure 4. Correlation Heat Map

understand patterns or groupings in the dataset. They can be interpreted to gain insights into the structure of the data. It helps discover hidden structures, identify similar groups, and facilitate further analysis or decision-making based on the identified clusters.

Figure 6 shows Precision, Recall, F1 Score, Accuracy, and Kappa Statistics for Logistic regression, SVM, Nave Bayes KNN, Grid Search Decision Trees, Random Forest, and XG Boost.

Table 4 shows the assessment of the accuracy percentage of our work with other proposed models. We

achieved an accuracy of 99.04%, better than many other proposed models. Our model compares with (Aldhyani et al., 2022; Fazlullah et al., 2023; Sagarkumar, 2023; GSR et al., 2023), which have an accuracy of 86.23%,96.53%, 92%, and 95%, respectively, as shown graphically in Figure 7 below. Table 5 shows the assessment of the precision percentage of our work with other proposed models. We achieved a precision of 95.06%, better than many other proposed models (Fazlullah et al., 2023; Emil et al., 2023), having 95.06% and 86.48%, respectively. Figure 8 shows the graphical representation of our model and other existing methods regarding precision percentage.

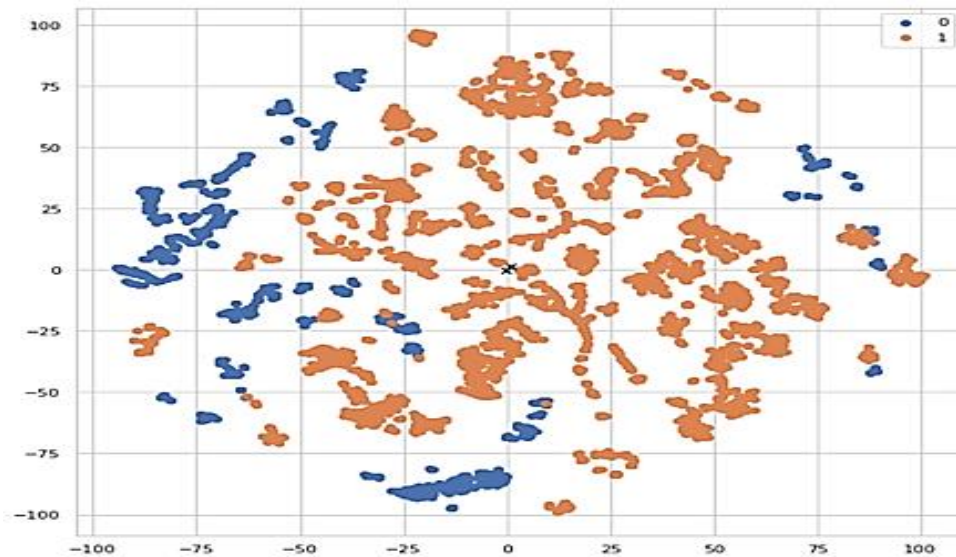


Figure 5. Cluster Analysis

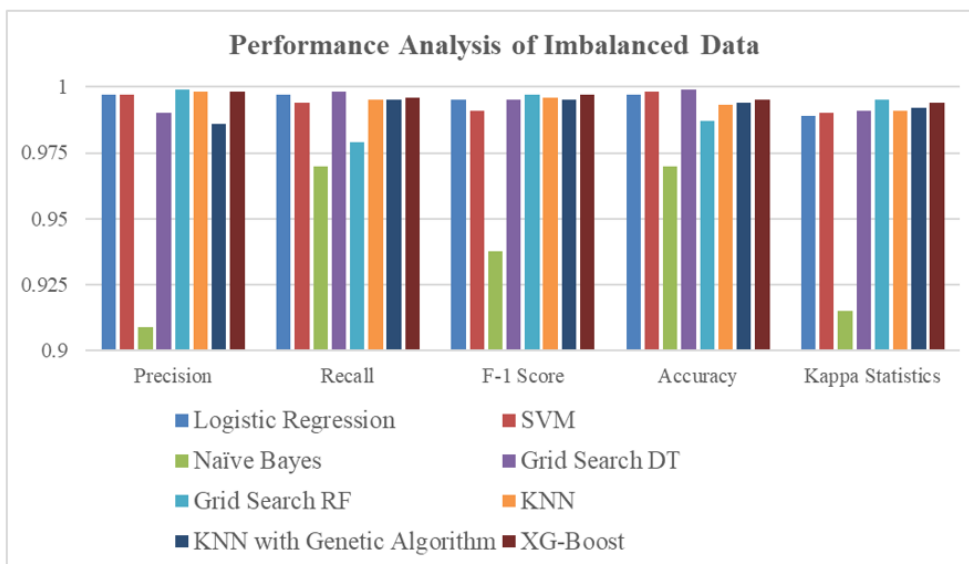


Figure 5. Performance Metrics on Oversamples Data

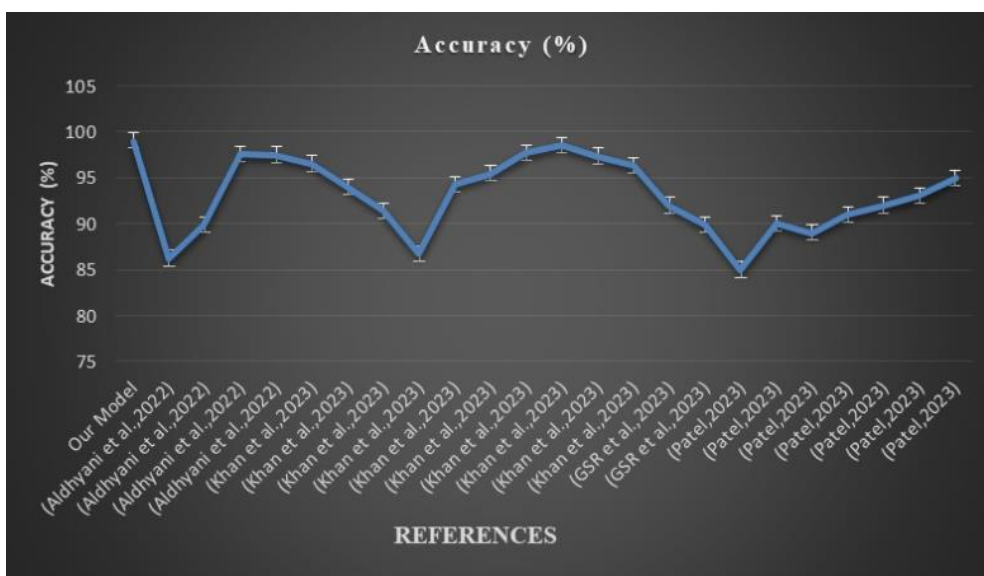


Figure 7. Graphical Analysis of Accuracy Model

Table 4. Performance Evaluation on Accuracy Parameter

Reference	Accuracy (%)
Our Model	99.04
Aldhyani et al., 2022	86.23
Aldhyani et al., 2022	89.84
Aldhyani et al., 2022	97.54
Aldhyani et al., 2022	97.50
Khan et al., 2023	96.53
Khan et al., 2023	94.05
Khan et al., 2023	91.41
Khan et al., 2023	86.72
Khan et al., 2023	94.32
Khan et al., 2023	95.46
Khan et al., 2023	97.69
Khan et al., 2023	98.56
Khan et al., 2023	97.37
Khan et al., 2023	96.33
GSR et al., 2023	92.00
GSR et al., 2023	89.89
Patel, 2023	85.00
Patel, 2023	90.00
Patel, 2023	89.00
Patel, 2023	91.00
Patel, 2023	92.00
Patel, 2023	93.00
Patel, 2023	95.00

Table 5. Model comparison in terms of Precision parameter

Reference	Precision (%)
Our Model	95.06
Khan et al., 2023	91.88
Khan et al., 2023	92.56
Khan et al., 2023	93.83
Khan et al., 2023	94.74
Khan et al., 2023	92.33
Khan et al., 2023	91.99
GSR et al., 2023	86.48
GSR et al., 2023	83.66

Conclusion

We present a way of detecting cloud attacks using a supervised learning technique and dataset. Our model gives 99.04 % accuracy, so in many practical scenarios, it can be used as discussed below: As cloud computing has emerged as new technological advancement and most businesses are deploying cloud services to boost their business, the cloud is becoming increasingly vulnerable to cryptographic attacks. These attacks can affect the smooth working of a business and can even lead to stilling relevant organizational information. Our model presents a supervised learning technique to detect cloud attacks with an accuracy of 99.04% and a precision of 95.06%. Classifiers like Logistic regression, simple vector machine (SVM), Random Forest, Decision Tree, Naïve Bayes, Xtreme Gradient Boosting (XGBoost), K-

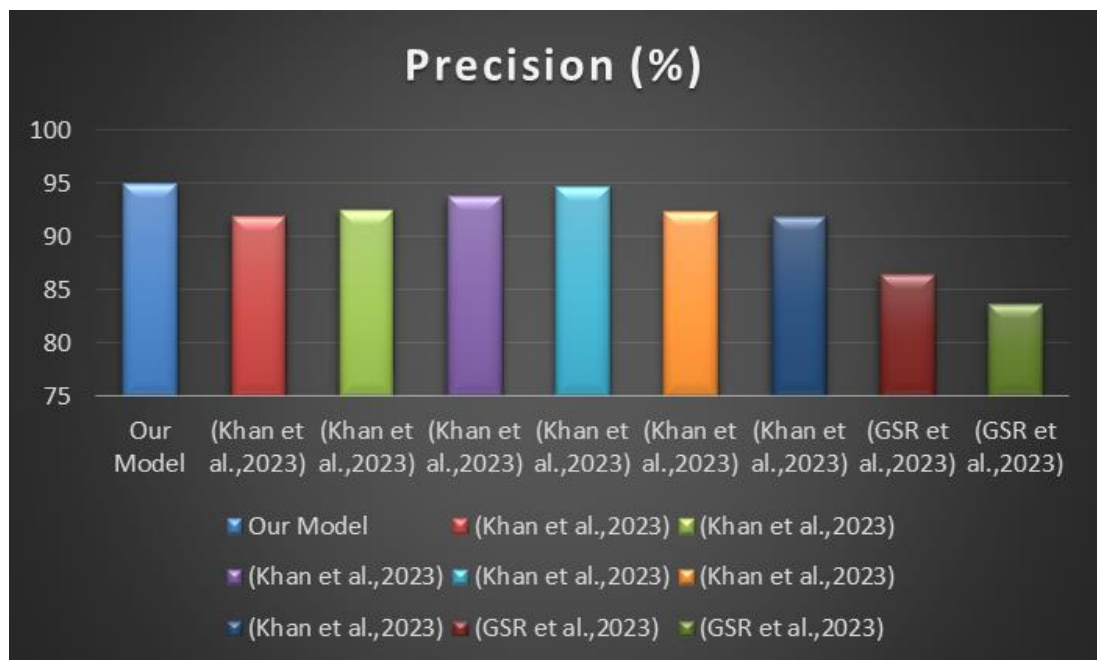


Figure 6. Comparison of Precision among various existing methods

Nearest Neighbour (K-NN), etc. are used in our experimental work. The model can prevent a cloud attack if deployed in the actual scenario. In the future, this model can be used to detect specific cloud attacks like Cross Site Scripting (XSS) and SQL Injection attacks.

Acknowledgment

The laboratory facilities and support for this study were provided by the Birla Institute of Technology, Mesra, Jharkhand, India, which the authors gratefully acknowledge.

Conflict of Interest

The authors declare no conflict of interest.

References

- Agrawal, N., & Tapaswi, S. (2019). Defense mechanisms against DDoS attacks in a cloud computing environment: State-of-the-art and research challenges. *IEEE Communications Surveys & Tutorials*, 21(4), 3769-3795. <https://doi.org/1109/COMST.2019.2934468>.
- Aldhyani, T. H. H., & Alkahtani, H. (2022). Artificial Intelligence Algorithm-Based Economic Denial of Sustainability Attack Detection Systems: *Cloud Computing Environments. Sensors*, 22(13), 4685. MDPI AG. Retrieved from <http://dx.doi.org/10.3390/s22134685>
- Anitha, P. T., Dibaba, W., & Boddu, R. (2023, May). Mitigation of Attacks Using Cybersecurity Deep Models in Cloud Servers. *IEEE, In 2023 International Conference on Disruptive Technologies (ICDT)*. pp. 202-205. <https://doi.org/10.1109/ICDT57929.2023.10150832>.
- Arunkumar, M., & Kumar, K. A. (2023). GOSVM: Gannet optimization based support vector machine for malicious attack detection in cloud environment. *International Journal of Information Technology*, 15(3), 1653-1660. <https://doi.org/10.1007/s41870-023-01192-z>
- Ashlam, A. A., Badii, A., & Stahl, F. (2023). Data-Mining and Hashing to Prevent Application-Layer DDoS and SQL Injection Attacks. In *2023 IEEE International Conference on Advanced Systems and Emergent Technologies (IC_ASET)*, pp. 01-06. https://doi.org/10.1109/IC_ASET58101.2023.10150694.
- Bag, S., Golder, R., Sarkar, S., & Maity, S. (2023). SENE: A novel manifold learning approach for distracted driving analysis with spatio-temporal and driver praxeological features. *Engineering Applications of Artificial Intelligence*, 123, 106332. <https://doi.org/10.1016/j.engappai.2023.106332>
- Bicego, M. (2023). DisRFC: a dissimilarity-based Random Forest Clustering approach. *Pattern Recognition*, 133, 109036. <https://doi.org/10.1016/j.patcog.2022.109036>
- Butt, U.A., Amin, R., Mehmood, M. (2023). Cloud Security Threats and Solutions: A Survey. *Wireless Pers Commun*, 128, 387-413. <https://doi.org/10.1007/s11277-022-09960-z>
- Chauhan, N., Kumar, V., & Dixit, S. (2023). To achieve sustainability in a supply chain with Digital integration: A TISM approach. *International Journal of Experimental Research and Review*, 30, 442-451. <https://doi.org/10.52756/ijerr.2023.v30.041>
- Clemens, V., Schulz, L. C., Gartner, M., & Hausheer, D. (2023, May). DDoS Detection in P4 Using Hyperloglog and Countmin Sketches. In *NOMS 2023-2023 IEEE/IFIP Network Operations and Management Symposium*, pp. 1-6. <https://doi.org/10.1109/NOMS56928.2023.10154315>.
- Das, S., & Sarkar, S. (2022). News media mining to explore speed-crash-traffic association during COVID-19. *Transportation Research Record*, 03611981221121261. <https://doi.org/10.1177/03611981221121261>
- Dash, G., Sharma, C., & Sharma, S. (2023). Sustainable Marketing and the Role of Social Media: An Experimental Study Using Natural Language Processing (NLP). *Sustainability*, 15(6), 5443. MDPI AG. Retrieved from <http://dx.doi.org/10.3390/su15065443>
- Dey, P., Chowdhury, S., Abadie, A., Yaroson, E. V., & Sarkar, S. (2023). Artificial Intelligence-Driven Supply Chain Resilience in Vietnamese Manufacturing Small-and Medium-Sized Enterprises. *International Journal of Production Research*. <https://doi.org/10.1080/00207543.2023.2179859>
- Emil Selvan, G. S. R., Ganeshan, R., Jingle, I., & Ananth, J. P. (2023). FACVO-DNFN: Deep learning-based feature fusion and Distributed Denial of Service attack detection in cloud computing. *Knowledge-Based Systems*, 261, 110132. <https://doi.org/10.1016/j.knosys.2022.110132>
- Gemmer, D. D., Meyer, B. H., de Mello, E. R., Schwarz, M., Wangham, M. S., & Nogueira, M. (2023, May). A Scalable Cyber Security Framework for the Experimentation of DDoS Attacks of Things.

- In NOMS 2023-2023 IEEE/IFIP *Network Operations and Management Symposium*, pp. 1-7.
<https://doi.org/10.1109/NOMS56928.2023.10154400>.
- George, A. S., & Sagayarajan, S. (2023). Securing Cloud Application Infrastructure: Understanding the Penetration Testing Challenges of IaaS, PaaS, and SaaS Environments. *Partners Universal International Research Journal*, 2(1), 24-34.
<https://doi.org/10.5281/zenodo.7723187>
- Gong, S., Ochiai, H., & Esaki, H. (2020). Scan-Based Self Anomaly Detection: Client-Side Mitigation of Channel-Based Man-in-the-Middle Attacks Against Wi-Fi. In 2020 *IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, pp. 1498-1503.
<https://doi.org/10.1109/COMPSAC48688.2020.00-43>.
- Gopinath, N., & Shyry, S. P. (2023). Side Channel Attack Free Quantum Key Distribution Using Entangled Fuzzy Logic. *Braz. J. Phys.*, 53, 35.
<https://doi.org/10.1007/s13538-022-01246-w>
- GSR, E. S., Ganeshan, R., Jingle, I. D. J., & Ananth, J. P. (2023). FACVO-DNFN: Deep learning-based feature fusion and Distributed Denial of Service attack detection in cloud computing. *Knowledge-Based Systems*, 261, 110132.
<https://doi.org/10.1016/j.knosys.2022.110132>
- Ha, G., Chen, H., Jia, C., & Li, M. (2022). Threat model and defense scheme for side-channel attacks in client-side deduplication. *Tsinghua Science and Technology*, 28(1), 1-12.
<https://doi.org/10.26599/TST.2021.9010071>.
- Iban, M. C., & Bilgilioglu, S.S. (2023). Snow avalanche susceptibility mapping using novel tree-based machine learning algorithms (XGBoost, NGBoost, and LightGBM) with eXplainable Artificial Intelligence (XAI) approach. *Stochastic Environmental Research and Risk Assessment*, 37(6), 2243-2270.
<https://doi.org/10.1007/s00477-023-02392-6>
- Jain, A., & Rajak, R. (2023). A systematic review of workflow scheduling techniques in a fog environment. *International Journal of Experimental Research and Review*, 30, 100-108.
<https://doi.org/10.52756/ijerr.2023.v30.011>
- Joshi, A., Capezza, S., Alhaji, A., & Chow, M. Y. (2023). Survey on AI and Machine Learning Techniques for Microgrid Energy Management Systems. *IEEE/CAA Journal of Automatica Sinica*, 10(7), 1513-1529.
<https://doi.org/10.1109/JAS.2023.123657>.
- Joshi, A., Capezza, S., Alhaji, A., & Chow, M. Y. (2023). Survey on AI and Machine Learning Techniques for Microgrid Energy Management Systems. *IEEE/CAA Journal of Automatica Sinica*, 10(7), 1513-1529.
<https://doi.org/10.1109/JAS.2023.123657>.
- Khan, F., Jan, M. A., Alturki, R., Alshehri, M. D., Shah, S. T., & ur Rehman, A. (2023). A Secure Ensemble Learning-Based Fog-Cloud Approach for Cyberattack Detection in IoMT. *IEEE Transactions on Industrial Informatics*, pp. 1-9.
<https://doi.org/10.1109/TII.2022.3231424>.
- Khurana, D., Koli, A., & Khatter, K. (2023). Natural language processing: state of the art, current trends and challenges. *Multimed. Tools Appl.*, 82, 3713-3744.
<https://doi.org/10.1007/s11042-022-13428-4>
- Kreuzberger, D., Kühl, N., & Hirschl, S. (2023). Machine learning operations (mlops): Overview, definition, and architecture. *IEEE Access*, 11, 31866-31879.
<https://doi.org/10.1109/ACCESS.2023.3262138>
- Kurani, A., Doshi, P., & Vakharia, A. (2023). A Comprehensive Comparative Study of Artificial Neural Network (ANN) and Support Vector Machines (SVM) on Stock Forecasting. *Ann. Data. Sci.*, 10, 183-208.
<https://doi.org/10.1007/s40745-021-00344-x>
- Kwekha-Rashid, A.S., Abduljabbar, H.N., & Alhayani, B. (2023). Coronavirus disease (COVID-19) cases analysis using machine-learning applications. *Appl. Nanosci.*, 13, 2013-2025.
<https://doi.org/10.1007/s13204-021-01868-7>
- Lu, Y., Qi, Y., Qi, S., Zhang, F., Wei, W., Yang, X., & Dong, X. (2021). Secure deduplication-based storage systems with resistance to side-channel attacks via fog computing. *IEEE Sensors Journal*, 22(18), 17529-17541.
<https://doi.org/10.1109/JSEN.2021.3052782>.
- Ma, T., Xu, C., Yang, S., Huang, Y., an, Q., Kuang, X., & Grieco, L. A. (2023). A Mutation-Enabled Proactive Defense against Service-Oriented Man-in-The-Middle Attack in Kubernetes. *IEEE Transactions on Computers*, pp. 1-14.
<https://doi.org/10.1109/TC.2023.3238125>
- Mohy-eddine, M., Guezzaz, A., Benkirane, S., & Azrour, M. (2023). An efficient network intrusion detection model for IoT security using K-NN classifier and feature selection. *Multimedia Tools and Applications*, pp. 1-19.

- <https://doi.org/10.1007/s11042-023-14795-2>
Paramanik, A. R., Sarkar, S., & Sarkar, B. (2022). OSWMI: An objective-subjective weighted method for minimizing inconsistency in multi-criteria decision making. *Computers & Industrial Engineering*, 169, 108138. <https://doi.org/10.1016/j.cie.2022.108138>
- Patel, S. K. (2022). Attack detection and mitigation scheme through novel authentication model enabled optimized neural network in smart healthcare. *Computer Methods in Biomechanics and Biomedical Engineering*, pp. 1-27. <https://doi.org/10.1080/10255842.2022.2045585>
- Patel, S.K. (2022). Attack detection and mitigation scheme through novel authentication model enabled optimized neural network in smart healthcare. *Computer Methods in Biomechanics and Biomedical Engineering*, pp. 1-27. <https://doi.org/10.1080/10255842.2022.2045585>
- Pramanik, A., Sarkar, S., & Maiti, J. (2021). A real-time video surveillance system for traffic pre-events detection. *Accident Analysis & Prevention*, 154, 106019. <https://doi.org/10.1016/j.aap.2021.106019>
- Radhakishan, V., & Selvakumar, S. (2011, September). Prevention of man-in-the-middle attacks using ID based signatures. *IEEE*, In 2011 Second International Conference on Networking and Distributed Computing, 165-169. <https://doi.org/10.1109/ICNDC.2011.40>
- Rajak, R., Choudhary, A., & Sajid, M. (2023). Load balancing techniques in cloud platform: A systematic study. *International Journal of Experimental Research and Review*, 30, 15-24. <https://doi.org/10.52756/ijerr.2023.v30.002>
- Ren, M., Tian, Y., Kong, S., Zhou, D., & Li, D. (2020, June). An detection algorithm for ARP man-in-the-middle attack based on data packet forwarding behavior characteristics. In 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC), *IEEE*, 1599-1604. <https://doi.org/10.1109/ITOEC49072.2020.9141555>
- Sahi, A., Lai, D., Li, Y., & Diyk, M. (2017). An efficient DDoS TCP flood attack detection and prevention system in a cloud environment. *IEEE Access*, 5, 6036-6048. <https://doi.org/10.1109/ACCESS.2017.2688460>
- Saleh, A., Dharshinni, N. P., Perangin-Angin, D., Azmi, F., & Sarif, M. I. (2023). Implementation of Recommendation Systems in Determining Learning Strategies Using the Naïve Bayes Classifier Algorithm. *Sinkron: Jurnal dan Penelitian Teknik Informatika*, 8(1), 256-267. <https://doi.org/10.33395/sinkron.v8i1.11954>
- Sarkar, S., Paramanik, A., Maiti, J., & Reniers, G. (2020). Predicting and analyzing injury severity: A machine learning-based approach using class-imbalanced proactive and reactive data. *Safety Science*, 125, 104616. <https://doi.org/10.1016/j.ssci.2020.104616>
- Sarkar, S., Vinay, S., Djeddi, C., & Maiti, J. (2021). Text mining-based association rule mining for incident analysis: a case study of a steel plant in India. In Pattern Recognition and Artificial Intelligence: 4th Mediterranean Conference, MedPRAI 2020, Hammamet, Tunisia, December 20–22, 2020, Proceedings Springer International Publishing, 4, 257-273. https://doi.org/10.1007/978-3-030-71804-6_19
- Sarkar, S., Vinay, S., Raj, R., Maiti, J., & Mitra, P. (2019). Application of optimized machine learning techniques for prediction of occupational accidents. *Computers & Operations Research*, 106, 210-224. <https://doi.org/10.1016/j.cor.2018.02.021>
- Sultan, A. B. M., Mehmood, S., & Zahid, H. (2022). Man in the Middle Attack Detection for MQTT based IoT devices using different Machine Learning Algorithms. *IEEE*, in 2022 2nd International Conference on Artificial Intelligence (ICAI), pp. 118-121. <https://doi.org/10.1109/ICCCNT54827.2022.9984365>
- Utukuru, S., Pisipati, R. K., & Karlapalem, K. (2023). Missing Data Resilient Ensemble Subspace Decision Tree Classifier. In Proceedings of the 6th Joint International Conference on Data Science & Management of Data (10th ACM IKDD CODS and 28th COMAD), pp. 104-107. <https://doi.org/10.1145/3570991.3571006>
- Verma, R., & Chandra, S. (2023). ReputE: A soft voting ensemble learning framework for reputation-based attack detection in fog-IoT milieu. *Engineering Applications of Artificial Intelligence*, 118, 105670. <https://doi.org/10.1016/j.engappai.2022.105670>
- Wang, N., Guo, H., Jing, Y., Zhang, Y., Sun, B., Pan, X., Chen, H., Xu, J., Wang, M., Chen, Xi, Song, L., & Cui, W. (2023). Development and validation of risk prediction models for large for gestational age infants using logistic regression and two

machine learning algorithms. *Journal of Diabetes*, 15(4), 338-348.

<https://doi.org/10.1111/1753-0407.13375>

Wu, K., Xu, Z., Lyu, X., & Ren, P. (2023). Cross-supervised learning for cloud detection. *GIScience & Remote Sensing*, 60(1), 2147298.

<https://doi.org/10.1080/15481603.2022.2147298>

Yan, Q., Yu, F. R., Gong, Q., & Li, J. (2015). Software-defined networking (SDN) and distributed denial of service (DDoS) attacks in cloud computing environments: A survey, some research issues, and challenges. *IEEE Communications Surveys & Tutorials*, 18(1), 602-622.

<https://doi.org/10.1109/COMST.2015.2487361>.

Yu, J., Yin, H., Xia, X., Chen, T., Li, J., & Huang, Z. (2023). Self-supervised learning for recommender systems: A survey. *IEEE Transactions on Knowledge and Data Engineering*,

<https://doi.org/10.1109/TKDE.2023.3282907>.

Zhang, Y., Mao, Y., Xu, M., Xu, F., & Zhong, S. (2019). Towards thwarting template side-channel attacks in secure cloud deduplications. *IEEE Transactions on Dependable and Secure Computing*, 18(3), 1008-1018. <https://doi.org/10.1109/TDSC.2019.2911502>.

How to cite this Article:

Animesh Kumar, Sandip Dutta and Prashant Pranav (2023). Supervised learning for Attack Detection in Cloud. *International Journal of Experimental Research and Review*, 31, 74-84.

DOI : <https://doi.org/10.52756/10.52756/ijerr.2023.v31spl.008>



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.