Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 4 (2025)
https://www.jneonatalsurg.com

Privilege Escalation Attack Detection And Mitigation In Cloud Using Machine Learning

Kande Vijaya Santhi[1], P. Neela Sundari[2]

[1]PG SCHOLAR, KKR&KSR Institute of Technology & Sciences, AP, India

[2]Assistant professor, KKR&KSR Institute of Technology & Sciences, AP, India

.Cite this paper as: Kande Vijaya Santhi, P. Neela Sundari, (2025) Privilege Escalation Attack Detection And Mitigation In Cloud Using Machine Learning, *Journal of Neonatal Surgery*, 14 (4), 536-541

ABSTRACT

The recent surge in the frequency and sophistication of cyber-attacks, coupled with the proliferation of smart devices, has posed significant cybersecurity challenges. While cloud computing has revolutionized modern business operations, its centralized architecture complicates the deployment of distributed security mechanisms, thereby increasing the risk of both accidental and malicious data breaches due to the vast amount of information exchanged between users and providers. Among these threats, malicious insiders with elevated access privileges pose a particularly severe risk. This study proposes a machine learning-based system to detect and classify insider threats by systematically identifying anomalous behaviors indicative of privilege escalation. To improve detection accuracy, ensemble learning techniques were employed across multiple algorithms, including Random Forest (RF), AdaBoost, XGBoost, and LightGBM, utilizing a customized dataset derived from the CERT insider threat dataset. Initial results indicated that LightGBM outperformed other models in overall accuracy. However, further experimentation revealed that Support Vector Machine (SVM) achieved the highest classification performance, particularly in identifying subtle insider behaviors. These findings suggest that a hybrid approach combining SVM with ensemble models could further enhance the robustness of insider threat detection systems.

Keywords: *Privilege Escalation, Insider Threat Detection, Cloud Security, Machine Learning, Ensemble Learning Algorithms.*

1. INTRODUCTION

Multiple studies have been presented regarding detecting irregularities and vulnerabilities in network systems to find security flaws or threats involving privilege escalation. But these studies lack the proper identification of the attacks. This study proposes and evaluates ensembles of Machine learning (ML) techniques in this context. They utilized the ‘‘CERT Insider Threat Tools’’ dataset since obtaining genuine business system logs is extremely challenging. Employee computer actions logs are included in the CERT dataset and certain organizational data such as employee’s departments and responsibilities. They built insider-threat detection models to emulate real world companies using machine learning-based methods. Privilege escalation attacks involve an attacker gaining higher-level access permissions than originally intended, potentially compromising the entire cloud infrastructure. Traditional security measures may not be sufficient to detect and prevent these sophisticated attacks. This research explores the integration of machine learning into cloud security to fortify defences against privilege escalation threats. To detect privilege escalation attempts, our system employs supervised machine learning models trained on historical data and anomaly detection algorithms. These models analyse patterns of user behaviour, system interactions, and access requests to identify deviations from normal activities. By continuously learning and adapting to evolving threat landscapes, the system can identify suspicious activities indicative of privilege escalation attempts.

2. LITERATURE REVIEW

Le et al. [9] discussed that insider threats are among the most expensive and difficult-to-detect forms of assault since insiders have access to a company’s networked systems and are familiar with its structure and security processes. A unique set of challenges face insider malware detection, such as extremely unbalanced data, limited ground truth, and behavioral drifts and shifts. Machine learning is used to analyze data at several levels of detail under realistic situations to identify harmful behaviors, especially malicious insider attacks. Random Forest beats the other ML methods, achieving good detection performance and F1-score with low false positive rates in most situations. The proposed work achieved an accuracy of 85% and a false positive rate of only 0.78%.

Janjua et al. [10] discussed that preventing malicious insiders from acting maliciously in an organization’s system is a significant cybersecurity challenge. The paper’s main goal is to use several Machine Learning approaches to classify email from the TWOS dataset. The following supervised learning techniques that have been used on the dataset are Adaboost, Naïve Bayes (NB), Logistic Regression (LR), KNN, Linear Regression (LR), and Support Vector Machine (SVM).

Experiments reveal that AdaBoost has the best classification accuracy for harmful and non-malicious emails, with a 98% accuracy rate. Although the model was trained on the original dataset, the data is limited. The model's results may be improved if the dataset is bigger.

Tripathy et al. [12] discussed that conventional web-based and cloud apps are vulnerable to the most popular online threats. One of the greatest threats to a SaaS application is the SQL injection attack. They construct and test the classification for SQL attack detection using machine learning methods. They explore the ability of machine learning models to identify SQL injection attacks, including the AdaBoost Classifier, Random Forest, and Deep Learning utilizing ANN, TensorFlow's Linear Classifier, and Boosted Trees Classifier. More important than malicious reading activities are malicious writing operations. The random forest classifier surpasses all others on the dataset and obtains better accuracy

Sun et al. [13] discussed that the network is becoming increasingly integral to businesses and organizations. So there is an increase in network security threats. Data leakage incidents from 15 nations and 17 industry groups were examined for Ponemon's 2018 Cost of a Data Breach Study, with 48% being malicious operations. While insiders' faulty actions were the cause of 27% of the incidents. They used the tree structure technique to study user behavior and create the feature sequence in this article. To distinguish between the feature patterns and detect unusual users, the COPOD approach is adopted. Additionally, the detection effect outperforms the standard unsupervised learning approach. Processing vast amounts of complicated and diverse data using this way provides benefits.

Liu et al. [15] discussed that information communication technology systems are increasingly vulnerable to cyber security attacks, most of which come from within the organization. Detecting and mitigating insider threats is a complicated challenge because insiders are hidden behind enterprise-level security defense measures and frequently have privileged network access. By gathering and reassembling information from the literature, they present the many types of insiders and the threats they bring. Insider threats are of three types: Masquerader, Traitor, and Unintentional perpetrator. Prevention may be viewed as a set of defensive procedures that can help prevent or enhance the identification of various internal threats. They examine the suggested efforts from a data analytics viewpoint, presenting them in terms of host, network, and contextual data analytics. Meanwhile, relevant studies are analyzed and compared, with a brief overview to show the benefits and drawbacks.

Abdelsalam et al. [24] discussed a Deep Learning-based malware detection technique (DL). Employing raw, process behavior (performance metrics) data, the study demonstrated the usefulness of using a 2D Convolutional Neural Network (CNN) for malware detection. The study illustrates the effectiveness of the proposed method by first developing a standard 2D CNN model which does not include the time window, and then making comparisons it to a newly developed 3D CNN model that greatly enhances detection accuracy, because of the use of a time window as the third dimension, thereby minimizing the problem of mislabeling. Results revealed a reasonable accuracy of 79% on the testing dataset by using 2D CNN.

3. METHODOLOGY

In recent exponential rise in attack frequency and sophistication, the proliferation of smart things has created significant cyber security challenges. Even though the tremendous changes cloud computing has brought to the business world, its centralization makes it challenging to use distributed services like security systems. Valuable data breaches might occur due to the high volume of data that moves between businesses and cloud service suppliers, both accidental and malicious. Attackers target data sources because they have the most valuable and sensitive information. Every cloud user's privacy and security are affected if data is lost. Insider threats are harmful operations carried out by people with authorization. With the fast growth of networks, many companies and organizations have established their internal networks. Malicious insiders become a crucial threat to the organization since they have more access and opportunity to produce significant damage. Unlike outsiders, insiders possess privileged and proper access to information and resources. Insider risks may be defined and addressed using criteria including insider indications, detection approaches, and insider kinds. There are two sorts of analysis intervals: real-time, which may identify malicious activity in real-time, and offline anomaly detection, which gathers log data and looks for certain patterns.

Dataset Selection & Preprocessing

The CERT Insider Threat Dataset is used as the primary data source for detecting privilege escalation attacks. This dataset contains logs of employee activities, including authentication attempts, file access, and system interactions. Data preprocessing involves cleaning, normalization, and feature extraction to ensure meaningful input for machine learning models. Key features such as user behavior patterns, access anomalies, and authentication logs are selected to enhance detection accuracy.

Machine Learning Model Implementation

Supervised learning techniques—including Random Forest (RF), AdaBoost, XGBoost, and LightGBM are employed to classify user activities as *normal* or *suspicious*. These models analyze user interactions such as login frequency, privilege usage, and resource access patterns.

To enhance detection robustness, ensemble learning strategies are applied, combining multiple model outputs to improve predictive accuracy.

For a majority voting ensemble, the final classification decision is given by**:**

$$\boldsymbol{H(x) = \mathrm{mode}\{h_1(x), h_2(x), \dots, h_n(x)\}}$$

For boosting methods like AdaBoost, the combined classifier is:

$$\boldsymbol{H(x) = \mathrm{sign}\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right)}$$

Where:

$h_t(x)$: individual weak learner at iteration

α_t : weight assigned to ht based on performance

These models are trained on a customized subset of the CERT insider threat dataset, ensuring relevance to real-world cloud behavior.

3.2 Performance Evaluation

The performance of each machine learning model is quantitatively assessed using standard classification metrics: accuracy, precision, recall, and F1-score. Let

TP, TN, FP, and FN represent the true positives, true negatives, false positives, and false negatives, respectively. The evaluation metrics are mathematically defined as follows:

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

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

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

Among the evaluated models, LightGBM achieves the highest overall detection accuracy, indicating its effectiveness in identifying insider threats across varied scenarios. However, Random Forest (RF) and AdaBoost demonstrate superior performance in detecting specific types of anomalies, such as behavioral biometrics deviations and abnormal command execution patterns. This suggests that while LightGBM is well-suited for general-purpose detection, RF and AdaBoost are valuable for specialized threat categories. Furthermore, Support Vector Machine (SVM), when included in the evaluation, surpassed all models in terms of overall precision, recall, and F1-score, highlighting its robustness and suitability for deployment in high-risk, real-world security environments.

3.3 Real-time and Offline Detection

The system operates in two complementary modes:

Real-time Detection: Continuously monitors user activity streams and triggers alerts for anomalous behavior.

Offline Detection: Periodically analyzes historical logs to identify long-term behavioral deviations.

To quantify anomaly, the following z-score–based anomaly detection is used:

$$\boldsymbol{\mathrm{Anomaly\ Score} = \frac{|x - \mu|}{\sigma}}$$

Where:

x = observed value

μ = mean (expected value)

σ = standard deviation

3.4 Integration with SIEM Systems

The ML-based detection system is integrated with Security Information and Event Management (SIEM) platforms, which centralize logs, correlate threat patterns, and provide actionable alerts.

Each detected anomaly is scored for risk prioritization using a weighted model:

$$\text{Threat Score}_i = w_1 \cdot A_i + w_2 \cdot C_i$$

Where:

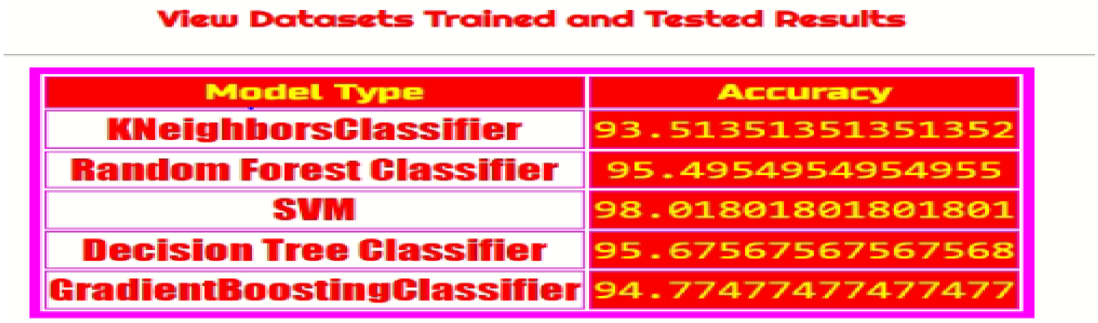
w1,w2 are weights,

Ai could represent anomaly score,

Ci could represent context score or confidence level.

This integration ensures that detected threats are not only identified but also contextually assessed, enabling swift incident response.

4. RESULTS AND DISCUSSIONS



View Datasets Trained and Tested Results

Model Type	Accuracy
KNeighborsClassifier	93.51351351351352
Random Forest Classifier	95.4954954954955
SVM	98.01801801801801
Decision Tree Classifier	95.67567567567568
GradientBoostingClassifier	94.77477477477477

Fig 1: Train & Test Results

Among the evaluated machine learning models, Support Vector Machine (SVM) exhibited the highest classification accuracy of 98.02%, making it the top-performing model in this study. This result is derived from a quantitative assessment using standard performance metrics, as visualized in both tabular and graphical formats. The significant accuracy achieved by SVM demonstrates its superior ability to distinguish between normal and anomalous behavior patterns within the insider threat dataset.

In the tabular results, SVM stands out by surpassing all other classifiers, including ensemble-based approaches such as Random Forest and Gradient Boosting. This underscores the model's effectiveness in handling high-dimensional behavioral data where decision boundaries may be non-linear and complex. The consistent precision in its performance output suggests that SVM can reliably reduce false positives and false negatives in sensitive security environments.

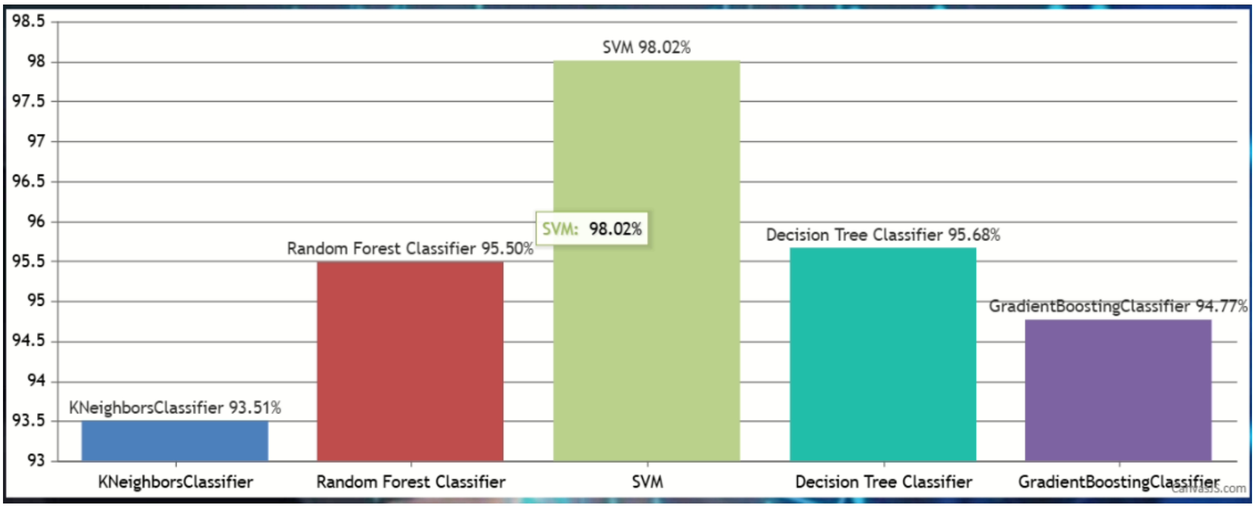


Fig 2: Visualization Report

The bar chart visualization further reinforces SVM's dominance, with its accuracy bar peaking visibly above all others at the 98.02% mark. This graphical evidence complements the numerical data, offering an intuitive and impactful representation of the model's effectiveness. Such high performance highlights SVM as a robust and dependable candidate for real-world insider threat detection systems, especially in scenarios requiring precise classification of user activities to prevent privilege escalation and other malicious behaviors.

5. CONCLUSION

Malicious insiders pose a significant threat to organizational security due to their authorized access and deeper knowledge of internal systems. Unlike external attackers, insiders can leverage their legitimate privileges to carry out stealthy and damaging attacks. This paper proposed a machine learning-based approach for detecting and classifying insider threats, utilizing a customized dataset derived from multiple files of the CERT insider threat dataset. Four ensemble-based supervised learning algorithms Random Forest (RF), AdaBoost, XGBoost, and LightGBM were applied, and their performances were evaluated using key classification metrics. Among these, LightGBM achieved an impressive accuracy of 97%. Further experimentation revealed that Support Vector Machine (SVM) achieved the highest classification accuracy of 98%, highlighting its potential as a leading model for insider threat detection. These findings demonstrate that machine learning models can significantly aid in early identification and mitigation of privilege escalation attacks, thus supporting secure cloud-based environments and enhancing organizational resilience against insider threats.

6. FUTURE SCOPE

Future enhancements include expanding the dataset for improved generalization, integrating deep learning techniques for better anomaly detection, enabling real-time threat identification, and adopting adaptive learning to handle evolving attacks. Additionally, incorporating explainable AI and integrating with advanced SIEM systems will enhance the robustness of insider threat detection in dynamic environments

REFERENCES

[1] U. A. Butt, R t, R. Amin, H. Aldabbas, S. Mohan, B. Alouffi, and A. Ahmadian, ‘‘Cloud-based email phishing attack using machine and deep learning algorithm,’’ Complex Intell. Syst., pp. 1–28, Jun. 2022.

[2] D. C. Le and A. N. Zincir-Heywood, ‘‘Machine learning based insider threat modelling and detection,’’ in Proc. IFIP/IEEE Symp. Integr. Netw. Service Manag. (IM), Apr. 2019, pp. 1–6.

[3] P. Oberoi, ‘‘Survey of various security attacks in clouds based environments,’’ Int. J. Adv. Res. Comput. Sci., vol. 8, no. 9, pp. 405–410, Sep. 2017.

[4] A. Ajmal, S. Ibrar, and R. Amin, ‘‘Cloud computing platform: Performance analysis of prominent cryptographic algorithms,’’ Concurrency Comput., Pract. Exper., vol. 34, no. 15, p. e6938, Jul. 2022.

[5] U. A. Butt, R. Amin, M. Mehmood, H. Aldabbas, M. T. Alharbi, and N. Albaqami, ‘‘Cloud security threats and solutions: A survey,’’ Wireless Pers. Commun., vol. 128, no. 1, pp. 387–413, Jan. 2023.

[6] H. Touqeer, S. Zaman, R. Amin, M. Hussain, F. Al-Turjman, and M. Bilal, ‘‘Smart home security: Challenges, issues and solutions at different IoT layers,’’ J. Supercomput., vol. 77, no. 12, pp. 14053–14089, Dec. 2021.

[7] S. Zou, H. Sun, G. Xu, and R. Quan, ‘‘Ensemble strategy for insider threat detection from user activity logs,’’ Comput., Mater. Continua, vol. 65, no. 2, pp. 1321–1334, 2020.

[8] G. Apruzzese, M. Colajanni, L. Ferretti, A. Guido, and M. Marchetti, ‘‘On the effectiveness of machine and deep learning for cyber security,’’ in Proc. 10th Int. Conf. Cyber Conflict (CyCon), May 2018, pp. 371–390.

[9] D. C. Le, N. Zincir-Heywood, and M. I. Heywood, ‘‘Analyzing data granularity levels for insider threat detection using machine learning,’’ IEEE Trans. Netw. Service Manag., vol. 17, no. 1, pp. 30–44, Mar. 2020.

[10] F. Janjua, A. Masood, H. Abbas, and I. Rashid, ‘‘Handling insider threat through supervised machine learning techniques,’’ Proc. Comput. Sci., vol. 177, pp. 64–71, Jan. 2020.

[11] R. Kumar, K. Sethi, N. Prajapati, R. R. Rout, and P. Bera, ‘‘Machine learning based malware detection in cloud environment using clustering approach,’’ in Proc. 11th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT), Jul. 2020, pp. 1–7.

[12] D. Tripathy, R. Gohil, and T. Halabi, ‘‘Detecting SQL injection attacks in cloud SaaS using machine learning,’’ in Proc. IEEE 6th Int. Conf. Big Data Secur. Cloud (BigDataSecurity), Int. Conf. High Perform. Smart Comput., (HPSC), IEEE Int. Conf. Intell. Data Secur. (IDS), May 2020, pp. 145–150.

[13] X. Sun, Y. Wang, and Z. Shi, ‘‘Insider threat detection using an unsupervised learning method: OPOD,’’ in Proc. Int. Conf. Commun., Inf. Syst. Comput. Eng. (CISCE), May 2021, pp. 749–754.

[14] J. Kim, M. Park, H. Kim, S. Cho, and P. Kang, ‘‘Insider threat detection based on user behavior modeling and anomaly detection algorithms,’’ Appl. Sci., vol. 9, no. 19, p. 4018, Sep. 2019.

[15] L. Liu, O. de Vel, Q.-L. Han, J. Zhang, and Y. Xiang, ‘‘Detecting and preventing cyber insider threats: A survey,’’ IEEE Commun. Surveys Tuts., vol. 20, no. 2, pp. 1397–1417, 2nd Quart., 2018.

[16] P. Chattopadhyay, L. Wang, and Y.-P. Tan, ‘‘Scenario-based insider threat detection from cyber activities,’’ IEEE Trans. Computat. Social Syst., vol. 5, no. 3, pp. 660–675, Sep. 2018.

[17] G. Ravikumar and M. Govindarasu, ‘‘Anomaly detection and mitigation for wide-area damping control using machine learning,’’ IEEE Trans. Smart Grid, early access, May 18, 2020, doi: 10.1109/TSG.2020.2995313.

[18] M. I. Tariq, N. A. Memon, S. Ahmed, S. Tayyaba, M. T. Mushtaq, N. A. Mian, M. Imran, and M. W. Ashraf, ‘‘A review of deep learning security and privacy defensive techniques,’’ Mobile Inf. Syst., vol. 2020, pp. 1–18, Apr. 2020.

[19] D. S. Berman, A. L. Buczak, J. S. Chavis, and C. L. Corbett, ‘‘A survey of deep learning methods for cyber security,’’ Information, vol. 10, no. 4, p. 122, 2019.

[20] N. T. Van and T. N. Thinh, ‘‘An anomaly-based network intrusion detection system using deep learning,’’ in Proc. Int. Conf. Syst. Sci. Eng. (ICSSE), 2017, pp. 210–214.

[21] G. Pang, C. Shen, L. Cao, and A. V. D. Hengel, ‘‘Deep learning for anomaly detection: A review,’’ ACM Comput. Surv., vol. 54, no. 2, pp. 1–38, Mar. 2021.

[22] A. Arora, A. Khanna, A. Rastogi, and A. Agarwal, ‘‘Cloud security ecosystem for data security and rivacy,’’ in Proc. 7th Int. Conf. Cloud Comput., Data Sci. Eng., Jan. 2017, pp. 288–292.

[23] L. Coppolino, S. D’Antonio, G. Mazzeo, and L. Romano, ‘‘Cloud security: Emerging threats and current solutions,’’ Comput. Electr. Eng., vol. 59, pp. 126–140, Apr. 2017.

[24] M. Abdelsalam, R. Krishnan, Y. Huang, and R. Sandhu, ‘‘Malware detection in cloud infrastructures using convolutional neural networks,’’ in Proc. IEEE 11th Int. Conf. Cloud Comput. (CLOUD), Jul. 2018, pp. 162–169.

[25] F. Jaafar, G. Nicolescu, and C. Richard, ‘‘A systematic approach for privilege escalation prevention,’’ in Proc. IEEE Int. Conf. Softw. Quality, Rel. Secur. Companion (QRS-C), Aug. 2016, pp. 101–108.

[26] N. Alhebaishi, L. Wang, S. Jajodia, and A. Singhal, ‘‘Modeling and mitigating the insider threat of remote administrators in clouds,’’ in Proc. IFIP Annu. Conf. Data Appl. Secur. Privacy. Bergamo, Italy: Springer, 2018, pp. 3–20.

[27] F. Yuan, Y. Cao, Y. Shang, Y. Liu, J. Tan, and B. Fang, ‘‘Insider threat detection with deep neural network,’’ in Proc. Int. Conf. Comput. Sci. Wuxi, China: Springer, 2018, pp. 43–54.

[28] I. A. Mohammed, ‘‘Cloud identity and access management—A model proposal,’’ Int. J. Innov. Eng. Res. Technol., vol. 6, no. 10, pp. 1–8, 2019.

[29] F. M. Okikiola, A. M. Mustapha, A. F. Akinsola, and M. A. Sokunbi, ‘‘A new framework for detecting insider attacks in cloud-based e-health care system,’’ in Proc. Int. Conf. Math., Comput. Eng. Comput. Sci. (ICMCECS), Mar. 2020, pp. 1–6.

[30] G. Li, S. X. Wu, S. Zhang, and Q. Li, ‘‘Neural networks-aided insider attack detection for the average consensus algorithm,’’ IEEE Access, vol. 8, pp. 51871–51883, 2020.

[31] A. R. Wani, Q. P. Rana, U. Saxena, and N. Pandey, ‘‘Analysis and detection of DDoS attacks on cloud computing environment using machine learning techniques,’’ in Proc. Amity Int. Conf. Artif. Intell. (AICAI), Feb. 2019, pp. 870–875.

[32] N. M. Sheykhkanloo and A. Hall, ‘‘Insider threat detection using supervised machine learning algorithms on an extremely imbalanced dataset,’’ Int. J. Cyber Warfare Terrorism, vol. 10, no. 2, pp. 1–26, Apr. 2020.

[33] M. Idhammad, K. Afdel, and M. Belouch, ‘‘Distributed intrusion detection system for cloud environments based on data mining techniques,’’ Proc. Comput. Sci., vol. 127, pp. 35–41, Jan. 2018.

[34] P. Kaur, R. Kumar, and M. Kumar, ‘‘A healthcare monitoring system using random forest and Internet of Things (IoT),’’ Multimedia Tools Appl., vol. 78, no. 14, pp. 19905–19916, 2019.

[35] J. L. Leevy, J. Hancock, R. Zuech, and T. M. Khoshgoftaar, ‘‘Detecting cybersecurity attacks using different network features with LightGBM and XGBoost learners,’’ in Proc. IEEE 2nd Int. Conf. Cognit. Mach. Intell. (CogMI), Oct. 2020, pp. 190–197.

[36] R. A. Alsowail and T. Al-Shehari, ‘‘Techniques and countermeasures for preventing insider threats,’’ PeerJ Comput. Sci., vol. 8, p. e938, Apr. 2022.

[37] Reddy, G. Raghupal, and G. Radha Devi. "Security Privacy Content and Impact of Trust in Social Networks.", IJAIST, vol 6, no 11, pp- 394-398, Nov 2017.