

Adaptive Insider Threat Detection in Cloud Platforms Using Ensemble Machine Learning Models

Syeda Arshiya Nausheen^{1*}, P. Pavani², S. Sindhuja², Ch. Adarsh², G. Bhavan Kumar²

¹Assistant Professor, ²UG Student, ^{1,2}Department of Computer Science and Engineering-Cyber Security,

^{1,2}Malla Reddy Engineering College and Management Sciences, Kistapur, Medchal, 501401, Telangana, India

*Corresponding author: syedaarshiya@mrem.ac.in

Received: 02.03.2025

Revised: 15.04.2025

Accepted: 25.05.2025

ABSTRACT

Insider attacks in cloud environments pose major security risks by compromising the confidentiality and integrity of critical data through unauthorized access escalation. Conventional defense mechanisms, such as manual monitoring and rule-based anomaly detection, are slow, inefficient, and incapable of identifying complex patterns within large-scale cloud activity logs. This research proposes an ensemble machine learning-based framework to detect insider privilege escalation by leveraging models such as Random Forest, AdaBoost, XGBoost, LightGBM, and CatBoost. These algorithms learn behavioral patterns from historical security events, enabling automated and adaptive detection of suspicious user activities with improved accuracy and reduced false positives. The system enhances scalability and responsiveness, supporting real-time threat identification while reducing dependence on manual intervention. Experimental analysis demonstrates that the ensemble learning approach outperforms traditional detection systems by effectively recognizing subtle privilege misuse behaviors, thereby strengthening cloud security and providing a proactive defense mechanism against evolving insider threats.

Key words: Cloud Security, Privilege Escalation, Ensemble Learning, Anomaly Detection, Cybersecurity

1. INTRODUCTION

Cloud computing has revolutionized data storage and processing by offering scalable, on-demand services. However, this centralization introduces security challenges, particularly concerning unauthorized data access and insider threats. [1] Insider threats, originating from within the organization, pose significant risks as they involve individuals misusing their authorized access to harm systems or steal data. These threats are especially concerning in applications like online social networks (OSNs), business platforms, and e-commerce systems, where sensitive user information and financial data are prevalent. [2] Detecting insider threats is challenging because these individuals have legitimate access to the company's systems and data, making traditional security tools like firewalls often ineffective against such threats.

To address latency and bandwidth issues associated with cloud computing, paradigms like fog and edge computing have emerged. Fog computing extends cloud services to the network's edge, processing data closer to its source, while edge computing involves data processing directly on devices [3].

In data access scenarios within OSNs, business platforms, and e-commerce applications, insider attacks can occur during routine operations [4]. For instance, employees or contractors with legitimate access might intentionally steal sensitive customer data or inadvertently expose the system to external threats through negligence. Insider threats are a serious danger for e-commerce businesses, as disgruntled

employees or ex-employees may attempt to steal company data or proprietary information, hoping to sell it to rival businesses. Mitigating these threats requires implementing strict access controls, conducting regular security training, and employing advanced intrusion detection systems. [5] Effective mitigation involves enforcing least privilege access, implementing strong authentication, continuous monitoring, and rapid incident response mechanisms.

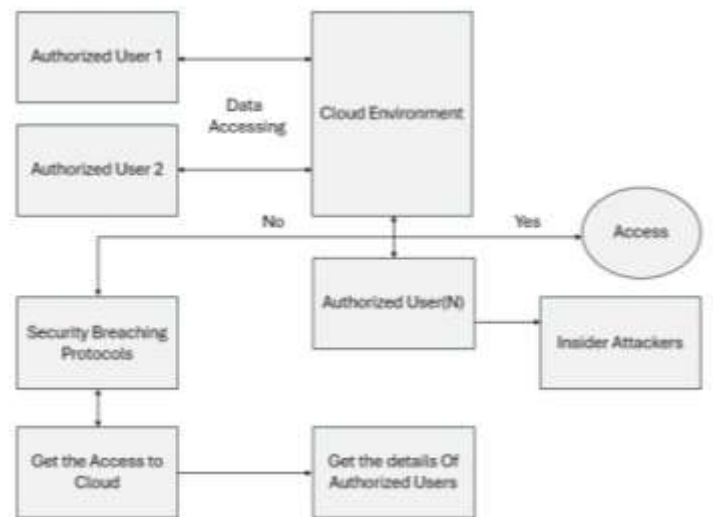


Fig 1. Cyber attacks Scenario.

Cloud environments provide seamless access to resources for multiple authorized users, but they are increasingly vulnerable to privilege insider attacks. Figure 1.1 shows the cyber-attacks detection scenario. Here, authorized users interact with the cloud, but malicious actors can exploit security weaknesses to gain unauthorized privileges. [6] Attackers use security breaching protocols to escalate their access, gaining control over sensitive cloud data. This rising threat necessitates robust detection and mitigation strategies to prevent unauthorized privilege escalation and secure cloud infrastructures.

One of the most concerning aspects of privilege insider attacks is insider threats, as depicted in the figure. [7] Attackers may exploit the credentials of authorized users or leverage insider access to retrieve sensitive details. This enables them to bypass security mechanisms and escalate privileges, gaining unrestricted access to critical cloud resources. [8] Traditional rule-based security measures are inadequate against such sophisticated attacks, emphasizing the need for an intelligent, adaptive approach that can identify and mitigate these threats in real time. Traditional attack detection systems in cloud computing environments often rely on security measures aligned with the Open Systems Interconnection (OSI) model, which delineates seven layers: physical, data link, network, transport, session, presentation, and application. [10] Each layer is fortified with specific protocols and defences to address potential vulnerabilities.

Attackers meticulously analyse these OSI layers to uncover weaknesses in frames, symbols, addresses, headers, and data payloads. By identifying and exploiting invalid packets and other anomalies, they can orchestrate attacks that bypass conventional security measures [11]. For instance, at the network layer, techniques like IP spoofing can be employed to masquerade as a trusted source, facilitating unauthorized access. Similarly, at the application layer, attackers might exploit vulnerabilities in software applications to gain control over a system [12].

These sophisticated methods often render basic protocols insufficient for comprehensive protection. Standard security measures may not detect or prevent attacks that exploit nuanced vulnerabilities across multiple layers. Consequently, there is a pressing need for advanced, multi-layered security approaches

that can dynamically analyze and respond to threats, ensuring robust protection in cloud computing environments[13].

To enhance security across all OSI layers, organizations are adopting advanced threat detection techniques. [14] These include behavioural analytics, which analyse user and entity behavior to detect deviations from normal patterns indicative of malicious activity, and the integration of threat intelligence feeds to stay informed about the latest threats and vulnerabilities. Additionally, deploying deception technologies, such as honeypots, can lure attackers, gather intelligence, and understand attack techniques.[15] Cloud providers also offer advanced threat detection tools, like AWS Guard Duty and Azure Advanced Threat Protection, which leverage machine learning and threat intelligence to provide comprehensive threat detection and response capabilities.

2. LITERATURE SURVEY

Najar, et al. [16] suggested a novel approach for low-rate DDoS detection and classification in cloud computing environments, addressing security threats that traditional methods struggle to mitigate. Proposed AE-CIAM, a hybrid framework integrating an attention-enhanced Autoencoder (AE) for feature extraction and a CNN Inception with Attention Mechanism for classification. Highlighted the model's effectiveness in balancing computational efficiency with high detection accuracy. Experimental evaluation on the CICIDS2017 dataset demonstrated superior performance, achieving 99.99% accuracy in binary classification and 99.44% in multiclassification, surpassing existing detection methods.

Sharma, et al. [17] proposed an integrated machine learning (ML) approach to enhance cloud computing security, addressing limitations of standalone ML models against emerging attack techniques. Suggested leveraging ensemble learning and AI techniques to improve prediction accuracy. Introduced a non-monotonous methodology within the ensemble model to enhance adaptive capabilities, facilitating the detection of unknown malicious activities and predicting zero-day attacks in cloud networks.

Nazari, et al. [18] This paper proposed an efficient cloud-based platform for real-time Cyber Threat Intelligence (CTI) detection, collection, and sharing from various online sources. Developed TSTEM, a prototype with a containerized microservice architecture leveraging Tweepy, Scrapy, Terraform, ELK, Kafka, and MLOps for autonomous IOC extraction and indexing. Implemented infrastructure as code (IaC) for provisioning and monitoring. Integrated custom focus-crawlers and multi-stage classifiers, enhanced by NLP models, to improve IOC identification. Experimental results showed high accuracy (exceeding 98%) in classification and extraction, ensuring rapid performance with minimal false positives, making it a robust solution for real-time CTI analysis and threat detection.

K V Kmet al. [19] This paper proposed a deep learning-based Intrusion Detection System (IDS) with a novel feature selection technique for cloud security. Focused on addressing security threats in cloud computing by overcoming limitations of traditional IDS. Designed a model with four phases: feature selection, pre-processing, classification, and encryption. Utilized the Adaptive Walrus Optimization Algorithm (AWO) for optimal feature selection, reducing computational complexity. Classified selected features using an Enhanced Adaptive Neuro-Fuzzy Inference System (EANFIS) for accurate intrusion detection.

Ahsun, et al. [20] This paper investigated the current state of real-time fraud detection, highlighting its limitations and emerging trends. Examined the role of artificial intelligence, blockchain, IoT, and other technologies in transforming fraud detection and prevention. Addressed challenges such as data privacy, regulatory compliance, and the evolving nature of fraud. Provided insights into enhancing real-time fraud detection to mitigate economic and social impacts. Discussed the challenges associated with implementing these technologies, including data privacy, regulatory compliance, and the evolving

nature of fraud. The findings of this study contributed to the understanding of how real-time fraud detection could be enhanced, ultimately reducing the economic and social impact of fraud.

Zhukabayeva, et al. [21] This paper explored cybersecurity challenges and solutions for cyber-physical systems (CPS) within IIoT and its integration with edge computing (IIoT–edge computing). We systematically analyzed literature from the past five years, identifying key IIoT layer attacks, intrusion methods, and threats affecting IIoT–edge environments. A detailed taxonomy of primary security mechanisms was developed, followed by a comparative analysis of our findings against existing research. The study highlighted vulnerabilities across the IIoT architecture, particularly concerning DoS, ransomware, malware, and MITM attacks. Advanced security technologies, including machine learning, federated learning, blockchain, deep learning, encryption, cryptography, IT/OT convergence, and digital twins, were examined as essential solutions for enhancing CPS security. The review underscored the need for real-time data protection and resilience in industrial operations. Finally, we outlined future research directions to strengthen cybersecurity in this evolving domain, ensuring robust protection against emerging threats.

Lytras, et al. [22] This paper Proposed key security concerns in cloud computing, including data loss, misconfiguration, unauthorized access, insecure APIs, account hijacking, and cyberattacks. Addressed data loss as a result of exposure due to breaches. Investigated misconfiguration as a major cause of data breaches, often due to reliance on CSPs for security. Highlighted unauthorized access risks due to cloud deployments being accessible from public networks. Focused on insecure APIs, which, if not secured, pose vulnerabilities. Provided insights into account hijacking, where stolen credentials grant access to sensitive data. Aimed to emphasize these threats and their implications for cloud security.

Singh, et al. [23] This paper Addressed the rapid advancement of cybersecurity threats over the past decade, with cybercriminals employing sophisticated techniques like ransomware, phishing, and advanced persistent threats (APTs). Highlighted the financial and reputational impacts of such attacks, emphasizing the need for robust cybersecurity measures. Investigated the role of machine learning in enhancing cybersecurity, leveraging large datasets to detect malicious patterns and anomalies. Focused on ML-driven improvements in real-time threat detection, response, and mitigation. Provided insights into intelligent machine learning solutions for cybersecurity while considering legal and ethical implications in a global context.

Srivastava, et al. [24] Reviewed recent developments in IoT technologies, highlighting security vulnerabilities in confidential data transmission. Addressed the role of Intrusion Detection Systems (IDS) in monitoring and alerting malicious activities. Investigated machine learning approaches for anomaly detection, noting the limited focus on sequential network data. Compared multiple models, including Random Forests (RF), Multi-Layer Perceptron's (MLP), and Long Short-Term Memory (LSTM), using the CIDDS-001 dataset. Provided experimental findings demonstrating that sequential detection is more effective than point-wise approaches. Highlighted LSTM's superior performance in detecting order-dependent traffic anomalies, achieving 99.94% accuracy and an F1-score of 91.66%.

3. PROPOSED SYSTEM

Cloud computing has revolutionized the way organizations store, manage, and process data by offering scalable and on-demand resources. However, with the increasing adoption of cloud services, security concerns such as privilege insider attacks have become a significant challenge. Privilege escalation occurs when an attacker exploits system vulnerabilities to gain unauthorized access to higher-level privileges, leading to severe security breaches, data theft, and system compromise. Detecting and mitigating such attacks in a cloud environment is crucial to ensuring data confidentiality, integrity, and availability. Traditional security mechanisms often fail to detect sophisticated privilege insider attacks

due to their evolving nature. Hence, machine learning (ML)-based detection systems provide an advanced and efficient approach to identifying and mitigating such threats.

In our proposed system, we focus on detecting and mitigating privilege insider attacks in cloud environments using machine learning. Specifically, we have chosen the Categorical Boosting Classifier (CatBoost) as our primary algorithm for attack detection. Unlike conventional security mechanisms or rule-based detection systems, ML algorithms can learn from historical attack patterns and identify anomalies in real time. While several ensemble learning algorithms, such as Random Forest, AdaBoost, XGBoost, and LightGBM, offer strong classification capabilities, we opted for CatBoost due to its unique advantages in handling categorical data, efficiency in training, and superior generalization on imbalanced datasets.

One of the key reasons for selecting CatBoost over other algorithms is its ability to handle categorical features without requiring extensive preprocessing such as one-hot encoding or label encoding. Since cloud security logs and access patterns often contain categorical attributes (such as user roles, access types, and privilege levels), CatBoost efficiently processes these features, reducing computational overhead. Moreover, CatBoost is designed to mitigate overfitting, ensuring better generalization on unseen attack patterns. Unlike XGBoost and LightGBM, which are highly sensitive to hyperparameter tuning, CatBoost provides robust performance with minimal tuning. Furthermore, its built-in feature importance mechanism enhances interpretability, allowing security analysts to understand which attributes contribute the most to attack detection.

The implementation of CatBoost in privilege escalation attack detection strengthens cloud security by offering an accurate, scalable, and interpretable solution. Our proposed system leverages real-time monitoring and ML-based classification to detect malicious privilege escalations before they can cause significant harm. By integrating CatBoost's powerful handling of categorical features, reduced preprocessing efforts, and superior generalization, we aim to build a highly efficient and intelligent security framework for cloud environments.

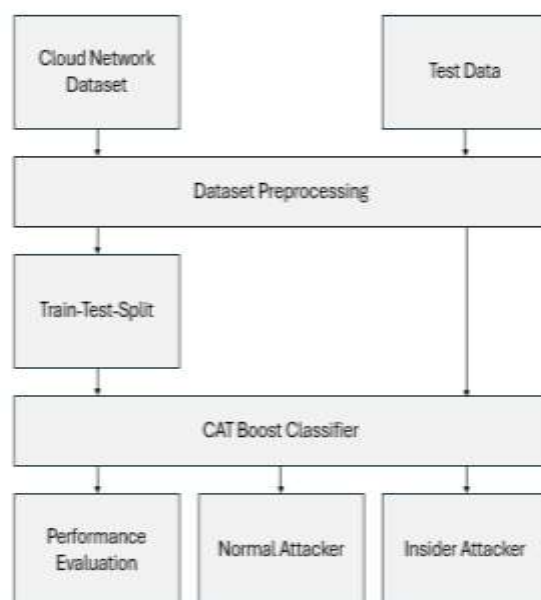


Figure 2. Proposed System Architecture.

The proposed system aims to detect and mitigate privilege insider attacks in cloud environments using machine learning (ML), specifically the CatBoost classifier. The architecture is designed to effectively

process and analyse cloud security logs to identify both normal users and insider attackers attempting to gain unauthorized privileges. The system begins by collecting data from a cloud network dataset, which includes user access logs, privilege levels, and activity patterns. This dataset is then subjected to a preprocessing step where we use the Standard Scaler to normalize numerical features, ensuring that different attributes contribute proportionally to the ML model's learning process. Following preprocessing, the dataset is split into training and testing sets using the train-test split method, allowing the model to learn attack patterns from historical data and validate its performance on unseen instances.

4. RESULT AND DESCRIPTION

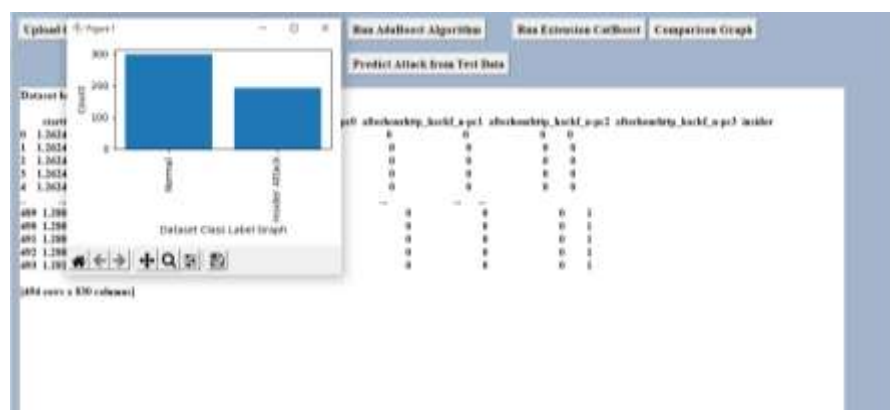


Fig 3. Load Dataset

Figure 3 shows the dataset upload function is triggered when a button is clicked to load and visualize a dataset. It opens a file dialog allowing the user to select a dataset file, which is then read into a Pandas DataFrame. The function clears any existing text in a text widget and displays a message indicating that the dataset has been successfully loaded. It then shows the contents of the dataset in the text widget. The function also analyzes the distribution of the 'insider' column, which indicates whether a record is classified as a "Normal" or "Insider Attack" instance. The distribution is visually represented as a bar graph, where the bars correspond to the count of each class, with the x-axis showing the class labels ("Normal" and "Insider Attack") and the y-axis displaying the frequency of each class. The graph is displayed in a window using Matplotlib, helping users easily observe the balance of class labels in the dataset.

Figure 4 shows preprocessing and train test splitting. The processDataset() function processes the loaded dataset by performing several key steps. First, it fills any missing values in the dataset with zeros using fillna(0). The dataset is then split into feature columns (X) and the target column (Y), with X containing all columns except the last one (which represents the target variable, insider). The feature columns (X) are normalized using StandardScaler to standardize the data, ensuring that each feature has a mean of 0 and a standard deviation of 1, which is crucial for many machine learning algorithms. Next, the dataset is shuffled using np.random.shuffle() to ensure that the data is randomly distributed before splitting. The dataset is divided into training and testing sets using train_test_split(), with 80% of the data used for training and 20% for testing. It ensures that the model is trained on the majority of the data while still being evaluated on a separate, unseen subset. Finally, the function displays the details of the dataset split (80% for training and 20% for testing) in the text widget, providing clear information on how the dataset has been divided for further model training and evaluation.

10.48047/jocaaa.2025.34.05.33

ensure that there are no NaN values, as the interfere with model predictions. The dataset is then converted into a numpy array with `data = dataset.values`.

- **Prediction:** The model (cat) is used to predict the class (either "Normal" or "Insider Attack") for each instance in the transformed dataset using `cat.predict(X)`, where X is the transformed input data.

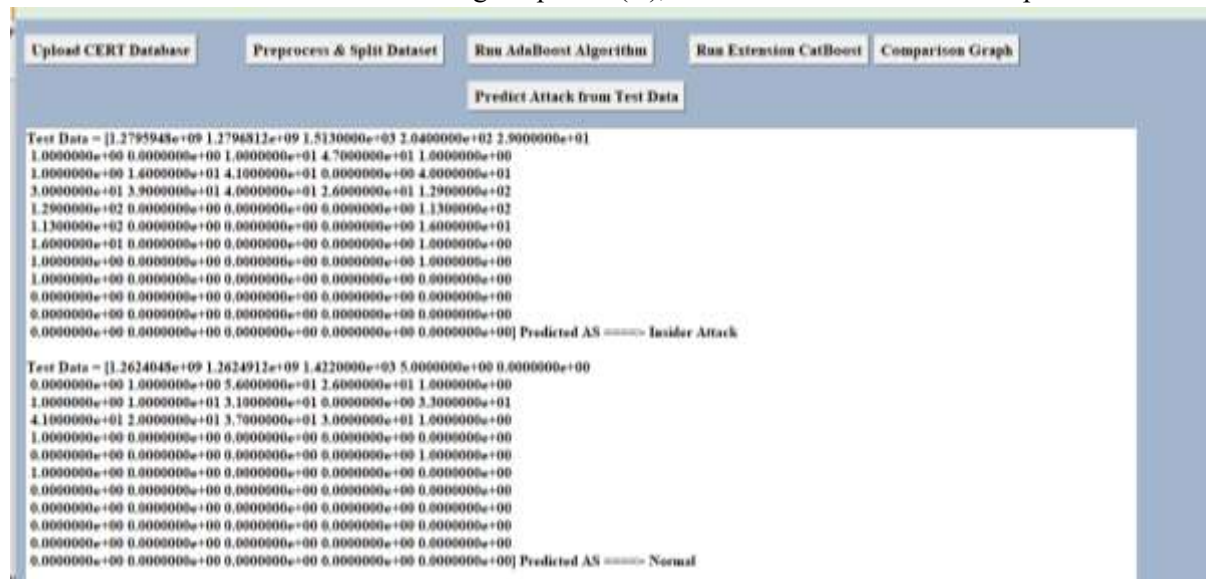


Fig 7. Predicted Outcomes.

Figure 7 shows predicted results from test data. The labels list contains the two possible classes: "Normal" and "Insider Attack". A loop iterates over the predictions. For each prediction (`predict[i]`), it converts the predicted value to an integer (`pred`), and based on the prediction value, the corresponding label (either "Normal" or "Insider Attack") is selected and printed. For each instance, the function inserts the input features (`data[i][0:60]`) and the predicted label into the output text box.

5. CONCLUSION

Privilege Escalation Attack Detection and Mitigation in Cloud Using Machine Learning" successfully demonstrates the potential of machine learning techniques in enhancing cloud security. The research addresses one of the critical challenges in cloud computing: insider attacks and privilege escalation, which often lead to severe data breaches and financial losses. By leveraging advanced algorithms such as AdaBoost and the enhanced CatBoost, the system efficiently detects malicious activities while minimizing false positives and false negatives.

Through rigorous experimentation and evaluation, the research highlights the importance of selecting the right machine learning model for effective threat detection. While traditional algorithms AdaBoost showed moderate accuracy, the CatBoost model emerged as the optimal choice due to its superior handling of categorical data and robust performance across diverse scenarios. The use of performance metrics such as accuracy, precision, recall, and F1-score ensures a comprehensive evaluation of the models' effectiveness.

Another significant achievement of the research lies in its ability to process and analyze real-time data for attack prediction. By preprocessing datasets and implementing scalable algorithms, the system provides an end-to-end solution that not only detects threats but also assists organizations in taking proactive measures to mitigate risks. The graphical representations and confusion matrices further make the system user-friendly and practical for stakeholders, including IT administrators and cybersecurity teams.

In conclusion, Research bridges the gap between existing security measures and the growing complexity of cyber threats in cloud environments. The successful implementation of the CatBoost algorithm

underscores the importance of adopting machine learning in security systems, paving the way for future advancements in cloud protection.

REFERENCES

- [1] S, A., D, S., & G, P. (2024). Malicious insider threat detection using variation of sampling methods for anomaly detection in cloud environment. In *Computers and Electrical Engineering* (Vol. 105, p. 108519). Elsevier BV.
- [2] Nascimento, E., Lins, L., Tavares, E., Pereira, P., Dantas, J., Kosta, S., & Maciel, P. (2025). Availability assessment of SDN-ICN service for multi-access edge computing. In *The Journal of Supercomputing* (Vol. 81, Issue 2). Springer Science and Business Media LLC.
- [3] Aleke, Ngozi Tracy, and Mohamed Trigui. "Legal and Ethical Challenges in Digital Forensics Investigations." In *Digital Forensics in the Age of AI*, pp. 147-176. IGI Global Scientific Publishing, 2025.
- [4] Rana, Sharif Uddin Ahmed. "Application of Data Mining Combined with K-means Clustering Algorithm in Enterprises' Risk Audit." In *The Economics of Talent Management and Human Capital*, pp. 29-54. IGI Global, 2025.
- [5] Madarie, R., Weulen Kranenbarg, M., & de Poot, C. (2025). Cybersecurity Expert Perspectives on Data Thieves' Actions in Digital Environments: Potential Refinements for Routine Activity Theory. In *Deviant Behavior* (pp. 1–19). Informa UK Limited.
- [6] Safeer, E. (2024). Reinforcement Learning Approaches in Cyber Security. In *Advances in Information Security, Privacy, and Ethics* (pp. 53–76). IGI Global.
- [7] Janjua, Faisal, Asif Masood, Haider Abbas, [10] Alzaabi, Fatima Rashed, and Abid Mehmood. "A review of recent advances, challenges, and opportunities in malicious insider threat detection using machine learning methods." *IEEE Access* 12 (2024): 30907-30927.
- [8] Leevy, J. L., Hancock, J., Zuech, R., & Khoshgoftaar, T. M. (2020). Detecting Cybersecurity Attacks Using Different Network Features with LightGBM and XGBoost Learners. In *2020 IEEE Second International Conference on Cognitive Machine Intelligence (CogMI)* (pp. 190–197). 2024 IEEE Second International Conference on Cognitive Machine Intelligence (CogMI).
- [9] Hunker, Jeffrey, and Christian W. Probst. "Insiders and Insider Threats-An Overview of Definitions and Mitigation Techniques." *J. Wirel. Mob. Networks Ubiquitous Computer Dependable Appl.* 2, no. 1 (2024): 4-27.
- [10] Asha, S., D. Shanmugapriya, and G. Padmavathi. "Malicious insider threat detection using variation of sampling methods for anomaly detection in cloud environment." *Computers and Electrical Engineering* 105 (2023): 108519.
- [11] Sivakumaran, A. R., G. Sreeja, Ch Poojitha, and I. Ramya. "Machine Learning for Cloud-Based Privilege Escalation Attack Detection and Mitigation with CATBOOST." *Turkish Journal of Computer and Mathematics Education* 15, no. 3 (2024): 164-176.
- [12] Inayat, Usman, Mashaim Farzan, Sajid Mahmood, Muhammad Fahad Zia, Shahid Hussain, and Fabiano Pallonetto. "Insider threat mitigation: Systematic literature review." *Ain Shams Engineering Journal* (2024): 103068.
- [13] Najjar, A. A., & Manohar Naik, S. (2024). AE-CIAM: a hybrid AI-enabled framework for low-rate DDoS attack detection in cloud computing. In *Cluster Computing* (Vol. 28, Issue 2). Springer Science and Business Media LLC.
- [14] Sharma, A., & Singh, U. K. (2025). Cloud Computing Security Through Detection & Mitigation of Zero-Day Attack Using Machine Learning Techniques. In *Natural Language Processing for Software Engineering* (pp. 357–388). Wiley.

10.48047/jocaaa.2025.34.05.33

- [15] Balasubramanian, P., Nazari, S., Kholgh, D. K., Mahmoodi, A., Seby, J., & Kostakos, P. (2025). A cognitive platform for collecting cyber threat intelligence and real-time detection using cloud computing. In *Decision Analytics Journal* (Vol. 14, p. 100545). Elsevier BV.
- [16] K V K, C., & V. , L. R. (2025). Automatic intrusion detection model with secure data storage on cloud using adaptive cyclic shift transposition with enhanced ANFIS classifier. In *Cyber Security and Applications* (Vol. 3, p. 100073). Elsevier BV.
- [17] Ahsun, Abbas, Beloved Joy, Berotian Noan, and Abiodun Okunola. "The Future of Real-Time Fraud Detection: Trends and Innovations." (2025).
- [18] Zhukabayeva, T., Zholshiyeva, L., Karabayev, N., Khan, S., & Alnazzawi, N. (2025). Cybersecurity Solutions for Industrial Internet of Things–Edge Computing Integration: Challenges, Threats, and Future Directions. In *Sensors* (Vol. 25, Issue 1, p. 213). MDPI AG.
- [19] Lytras, M. D., Alkhalidi, A. N., & Ordóñez de Pablos, P. (Eds.). (2024). Harnessing AI, Blockchain, and Cloud Computing for Enhanced e-Government Services. In *Advances in Electronic Government, Digital Divide, and Regional Development*. IGI Global.
- [20] Singh, B., & Kaunert, C. (2024). Intelligent Machine Learning Solutions for Cybersecurity. In *Advances in Computational Intelligence and Robotics* (pp. 359–386). IGI Global.
- [21] Srivastava, Rajeev, Harsh Raj, Shourya Dwivedi, Rohit Singh, and Nitish Chaurasiya. "Detection of phishing attacks using machine learning." In *Emerging Trends in Computer Science and Its Application*, pp. 105-109. CRC Press, 2025.
- [22] Sun, P., Wan, Y., Wu, Z., Fang, Z., & Li, Q. (2025). A survey on privacy and security issues in IoT-based environments: Technologies, protection measures and future directions. In *Computers & Security* (Vol. 148, p. 104097). Elsevier BV.