Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 4, April : 2025

THREAT IDENTIFICATION OF PRIVILEGE ESCALATION IN CLOUD USING ML

Mrs.B.Sowmya
Assistant professor
Usha Rama College Of Engineering And Technology
Telaprolu,AP, India
nissi.sowmya@gmail.com

Pallavi Kokkiligadda
UG Student in
Usha Rama College Of Engineering And Technology
Telaprolu,AP,India
pallavikokkiligadda22@gmail.com

Puvalla Gowtham
UG Student in
Usha Rama College Of Engineering And Technology
Telaprolu,AP,India
gowthampuvalla@gmail.com

Dondapati Srimanth
UG Student in
Usha Rama College Of Engineering And Technology
Telaprolu,AP, India
srimanthdondapati7@gmail.com

Shaik Tajuddin
UG Student in
Usha Rama College Of Engineering And Technology
Telaprolu,AP, India
tajuddinshaik2003@gmail.com

Abstract— ***With the widespread adoption of smart devices, cybersecurity is facing major challenges, particularly with the increasing frequency and sophistication of cyberattacks. Business operations are now much more streamlined thanks to cloud computing, but its centralized nature makes it harder for hackers to implement distributed security measures, leading to an increased risk of data breaches. Among these threats, the privileged access of malicious insiders is critical to making them cause severe harm. Such threats are often not easily detectable through traditional security measures.***

The research suggests a system that utilizes machine learning to identify and categorize insider threats, with specialized attention given to privilege escalation attacks. The system employs ensemble learning techniques to improve the predictive power of machine learning models. Using a custom dataset obtained from various versions of the CERT dataset, this research evaluates four machine learning algorithms: Random Forest (RF), Adaboost, XGBoost, and LightGBM. All five algorithms are evaluated. RF, Adaboost, and XGBoost are the other models that achieve higher accuracy, with LightGBM scoring the highest at 97%. The results were also impressive.

However, some algorithms are more efficient in identifying particular kinds of insider attacks, such as behavioral biometrics attacks.' Hence, it is possible to combine more than one model into a single classification system. This study offers valuable lessons in identifying and detecting insider threats, as well as improving cybersecurity measures within cloud-based organizational networks.

Keywords:
Cybersecurity, Insider Threat Detection, Privilege Escalation, Machine Learning, Ensemble Learning, Cloud Security, Data Breaches, Malicious Insiders, Smart Devices, Random Forest (RF), Adaboost, XGBoost, LightGBM, Behavioral Biometrics Attacks, CERT Dataset, Anomaly Detection, Threat Classification.

1.INTRODUCTION

Industry productivity and connectivity have been transformed by the rapid proliferation of smart devices and cloud computing. Moreover, Even so, this shift towards digital has posed significant security concerns. Organizations are facing more sophisticated cyberattacks that aim at critical data and infrastructure. The privileged access of insiders makes them a significant threat to the security of interiors, which can be exploited for malicious purposes

The centralized nature of cloud computing makes it difficult to implement distributed security measures, despite its many benefits. Businesses and cloud service providers exchange sensitive data on a regular basis, which can lead to both intentional or unintentional breaches. Organizations may be at risk from insider threats, as traditional security measures often lack the ability to detect and mitigate them.

External cyber threats are triggered by internal sources, while insider threats arise from within an organization. These threats may be intentional, such as when an employee intentionally exploits access for personal or financial gain, but may also be unintentional, like when security is compromised due to human errors or negligence. Because they have permission from insiders, anomalous behavior is difficult to detect in such systems.

A highly dangerous form of insider attack is the increase in privileges.'... An individual is granted unauthorized access to confidential information or administrative privileges in this kind of attack.' Various techniques, such as software vulnerabilities and misconfigurations, can lead to an increase in privileges. Such attacks can be very serious, with data stolen, losing money and damaging reputations if not detected early enough.

With the rise of machine learning, cybersecurity now finds it particularly advantageous as it provides automated and intelligent threat detection capabilities. Machine learning models can detect patterns of malicious activity by analyzing large data sets, unlike traditional security measures that rely

on rule-based methods. Insider threats can be detected with machine learning, which is particularly adept at detecting behavioral patterns.

By utilizing various machine learning models together, ensemble learning has been proven to be an effective method for identifying threats. Ensemble methods can be used to increase the predictive power of ensembles and reduce false positives by incorporating various strengths of each model. In this research, researchers explore the effectiveness of ensemble learning in identifying insider threats, with a focus on increasing privilege levels.

The study examines four machine learning algorithms, namely Random Forest (RF), Adaboost, the cloud computing technology known as XGBoost, and LightGBM. Among the algorithms, each offers unique advantages in classification tasks, making them useful for uncovering insider threats. The study's objective is to ascertain which model can accurately identify different types of insider attacks. Why?

A specific dataset from the CERT dataset is utilized to train and test machine learning models. This study is based on the CERT dataset, which contains numerous simulated insider threats that are highly respected and widely recognized. Through the use of multiple files from this dataset, the study provides a comprehensive and varied assessment of insider attack detection methods.

LightGBM is the more accurate model, with an accuracy rate of 97% (see first results).". It is evident that LightGBM has the ability to detect insider threats. In contrast, RF and Adaboost models exhibit superior performance in identifying particular types of insider threats, such as those that involve behavioral biometrics.

The results indicate that a hybrid approach with multiple machine learning models is required to improve classification accuracy. The integration of multiple algorithms could enable the development of a more comprehensive detection system that can identify broader threats within organizations.' This tactic has the potential to greatly enhance cybersecurity safeguards in cloud computing.

One of the primary obstacles to identifying insider threats is the significant variation in user behavior. Some unusual activities may not be a sign of unauthorized access; however, learning models is essential to differentiate between ordinary behavior and authentic security risks. Finding this balance is essential for reducing false alarms and increasing detection efficiency. Why?

Additionally, it is crucial to detect insider threats in real-time to prevent attacks before they become more severe. Machine learning models must be able to process large datasets efficiently and generate alerts quickly when suspicious activity is detected, which is necessary. Proper security measures require the use of real-time detection mechanisms.

Machine learning models' interpretation is a crucial aspect to consider. Understanding the reasoning behind a model's classification of an activity as threatening behavior is crucial for security analysts. Effective response strategies can be impeded by the limited visibility of black-box models, which offer only partial insight into their decision-making process. As a result, explainable AI methods must be included to increase trust and transparency in security solutions that incorporate machine learning.'

This study forms one part of a larger effort to improve cybersecurity in cloud-based environments. The research demonstrates that security organizations can enhance their ability to detect and mitigate risks by utilizing ensemble learning. Machine learning techniques are now capable of significantly lowering the likelihood of privilege escalation attacks and other insider threats.

The future work could involve exploring further machine learning techniques, such as deep learning and reinforcement learning, to enhance threat identification abilities. Additionally, combining threat intelligence data with behavioral analytics could improve the accuracy and responsiveness of detection systems. The preservation of cybersecurity threats requires ongoing research in this area.

Finally, insider threats remain a critical problem for organizations that operate in cloud environments.' Inadequate machine learning techniques are frequently required to detect and mitigate threats, which is why traditional security measures are insufficient. Organizations can use ensemble learning to construct more efficient and adaptable insider threat detection systems.

The outcomes of this research highlight the need for integrating various machine learning algorithms to achieve accurate threat classification. The need for intelligent security measures to safeguard sensitive data and secure cloud computing infrastructures will arise due to evolving cyber threats.

II LITERATURE REVIEW

The sophistication of cyber threats, particularly insider attacks, has led to a significant increase in cybersecurity concerns. A range of techniques have been developed by researchers to identify and counter insider threats, with machine learning being a recent area of focus. Within this section, there are published works that examine the detection of insider threats, such as privilege escalation attacks, cloud security, and machine learning-based techniques.

Many studies have indicated that insider risks in organizational networks are becoming more prevalent. The difficulty in identifying insider attacks is attributed to the fact that they have legitimate access, as stated by Bishop et al. [1]. According to the study, there is a need for anomaly detection mechanisms that can detect patterns of user behavior. Greitzer et al. [2] point to psychological and behavioral indicators that could be used to predict insider threats.

Insider threat detection has been made more accurate by the use of machine learning. A machine learning framework that utilizes user activity logs to detect suspicious behavior is suggested by Liu et al. [3].

The identification of insider threats is made possible by the use of supervised learning techniques, such as decision trees and support vector machines (SVM), according to their findings. The models' generalization difficulty often leads to a high rate of false-positives. In cybersecurity, classification performance can be improved through the implementation of

ensemble learning. Xu et al. [4] demonstrate the use of ensemble methods like Random Forest and Gradient Boosting Machines for anomaly detection in cloud environments. The accuracy of detection is significantly enhanced by using multiple classifiers together, as per their findings.

Several studies have examined privilege escalation attacks, which are a particular type of insider threat. The. The use of weak authentication mechanisms and misconfigurations by attackers is frequently exploited to gain unauthorized privileges, as noted by Hu et al. [5]. Their research indicates that early detection can be aided by scrutinizing system logs and access patterns. The concept of an access control mechanism that dynamically adjusts user privileges based on risk assessment models is also proposed by [6] Yin and al.

Researchers have widely incorporated the CERT insider threat dataset to assess insideer attack detection methods. This dataset is utilized by Mountrouidou et al. [7] to construct a deep learning-based detection system. They have a model that is highly accurate, but it requires incredibly much computation power for real time detection to be successful. Rather than other methods, Ahmed et al. [8] use feature selection techniques to optimize machine learning models while maintaining accuracy in detection precision as an alternative approach.

Several studies have been conducted on the use of behavioral analytics to detect insider threats. The approach of Eberle et al. [9] involves using graphs to model user interactions and identify anomalies by measuring changes in normal behavior.[a]. Although their method is highly accurate, it demands significant computational resources. Legg et al. have also examined the use of behavioral biometrics in insider threat detection, and found that keystroke dynamics and mouse movement patterns are reliable indicators of malicious intent.

Insider threat research has also delved into cloud security. The security risks of cloud environments are explored by Wang et al. [11] through the use of machine learning and access control policies as the foundation for developing a framework. According to their research, insider threats can be effectively reduced through the implementation of real- time monitoring. In the same vein, Takabi et al. [12] highlight the difficulties associated with implementing security policies in multi-tenant cloud environments and recommend adaptive authentication mechanisms as an alternative.

The use of machine learning has contributed to the improvement in insider threat detection capabilities. Zhang et al. [13] demonstrate the use of deep reinforcement learning in cybersecurity, emphasizing its ability to detect adaptive threats. Changing attack patterns lead to dynamical changes in detection thresholds in their model. In the meantime, Shabtai et al. [14] suggest an unsupervised learning technique for detecting anomalies in cloud systems, which decreases dependence on labeled datasets.

In spite of these advancements, uncovering insider threats remains a formidable undertaking. The high false-positive rates of various existing approaches have led to security analysts becoming alerted due to alert fatigue. The difficulty of interpreting machine learning models persists, particularly when considering the complexity of decision-making processes in black-box models.

In order to overcome these issues, this study suggests an ensemble learning-based approach that combines various machine learning algorithms for better classification accuracy. The study aims to utilize a proprietary CERT dataset to create an effective insider threat detection system that can identify and target privilege escalation attacks with significant accuracy.

These studies that have been reviewed offer valuable information on a variety of ways to identify insider threats. There are still gaps in the classification of privilege escalation attacks and in merging different models to improve detection performance. By utilizing various machine learning algorithms such as Random Forest, AdaBoost (for Deep Learning), and XGBoost (3Division Packet Topography), this research expands on earlier work to identify threats from within the system.

Finally, insider threat detection is an important aspect of cybersecurity and especially in cloud environments where sensitive data could be hacked. Ensemble methods have been a valuable addition to machine learning in terms of improving detection efficiency.

III. PROPOSED SYSTEM

A new insider threat detection system using machine learning-based techniques is proposed in this study to help counter the growing risks of inside attack, including privilege escalation attacks in cloud environments. The system's objective is to detect irregularities in the organization' LAN and identify insider threats by utilizing ensemble learning techniques. The system proposed involves the use of various machine learning algorithms with the aim to reduce false positives and improve accuracy in detecting malicious behavior.

The fundamental aim of this mechanism is to detect insider threats in a systematic manner, with sensitivity to those that arise from privilege increases. The privilege of users to access critical systems and data has made them a significant risk when their credentials are misused or compromised. Why is this so? The detection of such threats is difficult even for traditional security measures like rule-based systems. As a result, the proposed system uses machine learning to improve detection capabilities.

A unique dataset is employed by this system, which is derived from various versions of the CERT dataset. The CERT dataset is frequently utilized for insider threat detection research, as it contains simulations of user activities, including both benign and malicious behavior. The system preprocesses data to identify relevant features and helps machine learning models differentiate between normal and suspicious activity.

There are four machine learning algorithms that run on top of them: Random Forest (RF), Adaboost, XGBoost, and LightGBM. Despite their limitations, each algorithm has its own set of capabilities for classification tasks. In cases where the data size is large, Random Forest has a strong performance record and avoids overfitting. By examining

misclassified instances, Adaboost enhances classification accuracy. The combination of speed and performance in XGBoost makes it ideal for large-scale data analysis. The study's most accurate detection was achieved through LightGBM, a software that is optimized for high efficiency and scalability, making it ideal for real-time threat detection.

This proposed system would be based on ensemble learning. The reduction of false positives and improved predictive performance can be achieved through ensemble learning, which involves merging multiple models. This system examines a range of ensemble methods, such as voting and stacking, to integrate the strengths of individual models. A combination of techniques is employed by this system, which can identify a broad spectrum of insider threats with precision.

Feature selection is an essential aspect of the proposed system's success. Not all of these features in the dataset are equally useful for identifying threats. Consequently, feature engineering approaches the system to identify the most relevant features that indicate attempts to increase privileges. The features comprise login attempts, file access patterns, privilege changes, and atypical system executions.etn (user log in events).

The system is proactive threat detection and operates in real-time.?... Traditional security measures are limited in their ability to prevent attacks because they rely heavily on post-incident analysis. Through real-time data processing, the proposed system can generate alerts when suspicious activities are identified, enabling security teams to react quickly.

Diving between legitimate user actions and malicious intent is a common issue in insider threat detection. False positives can be generated by mimicking attack patterns in regular user behavior. By utilizing behavioral analytics and anomaly detection methods, the system proposed resolves this issue. Through continuous investigation of user behavior, the system can learn about new threats and decrease false alarms.

It also emphasizes interpretability. The reason for flagging a specific action as threatening is crucial for security analysts. Detailed AI methods are implemented to provide guidance on the model's decision-making process, which is achieved through this approach. The system's ability to provide detailed explanations for predictions helps security teams maintain trustworthiness and usability.

A second important consideration is scale. The system must be proficient in handling the large datasets of security data that organizations generate. This is a significant challenge. LightGBM's speed and memory usage optimization ensure that the system can scale without compromising performance. Moreover, cloud-based deployment alternatives are examined to ensure easy integration into enterprise environments.

Extensive experiments are conducted on the CERT dataset to verify the efficacy of the proposed system. Training and testing sets are included in the dataset, and a variety of performance metrics such as accuracy, precision, recall, F1-score, etc. are scrutinized. LightGBM is the most accurate algorithm in terms of classification, with a 97% accuracy, as demonstrated by the results. Nevertheless, Random Forest and Adaboost's proficiency in identifying particular kinds of insider dangers raises the possibility for a hybrid setup.

Insider threat detection methods are utilized to evaluate the system's effectiveness. The high false-positive rates and limited adaptability present a challenge for both traditional rule-based systems and standalone machine learning models. However, there are exceptions to this approach. The proposed system, which incorporates ensemble learning and real-time analytics, offers enhanced detection capabilities.

The proposed system also addresses deployment considerations. An organization's security needs to be well-prepared for the integration of machine learning-based threat detection. Designed to be compatible with current Security Information and Event Management (SIEM) solutions, the system provides easy access control over insider threats and quick response capabilities.

In addition, the proposed system incorporates features for ongoing learning and adaptation. The constant evolution of cyber threats means that static models may lose their effectiveness over time. In order to cope with this, the system employs incremental learning techniques that ensure models are updated as new threat patterns arise.

There are also ethical considerations. Although insider threat detection is critical for organizational security, user privacy must be safeguarded and surveillance should not be excessive. The proposed system is ethical as it concentrates on behavioral anomalies rather than storing personal data, thus meeting the requirements for privacy.[Note missing].

Finally, the system proposed is a powerful and efficient way to detect insider threats in cloud environments. The system incorporates machine learning, ensemble searching, real-time analytics and explainable AI to improve cybersecurity protection against privilege escalation attacks. Enhancements to the detection accuracy may involve the use of deep learning models and the refinement of feature selection techniques in future versions.

IV. WORK FLOW

The proposed insider threat detection system's workflow is structured to efficiently detect and classify privilege escalation attacks. It starts with data entry, then features are extracted, model training is conducted, evaluation performed, and finally real-time threat detection is performed.[Note 1]. All stages have been developed with the aim to improve both the efficiency of machine learning models and also provide security for cloud environments.

Data collection is the first step in this workflow; security logs and user activity data are gathered by the system from multiple sources. CERT insider threat dataset is the main data source for training and testing purposes. This dataset encompasses a range of user activities, from routine actions to harmful ones, such as privilege attempts. The data that is gathered includes login information, file access logs, command executions and change in privileges.

The process starts with the collection of data. During this stage, the data is cleaned up by taking care of any missing

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 4, April : 2025

numbers, eliminating duplicates, and normalizing numerical values.mean. The encoded information, such as access levels and user roles, is then used by machine learning models. Anomalous data entries are also analyzed to ensure that the dataset accurately represents both normal and malicious activities.

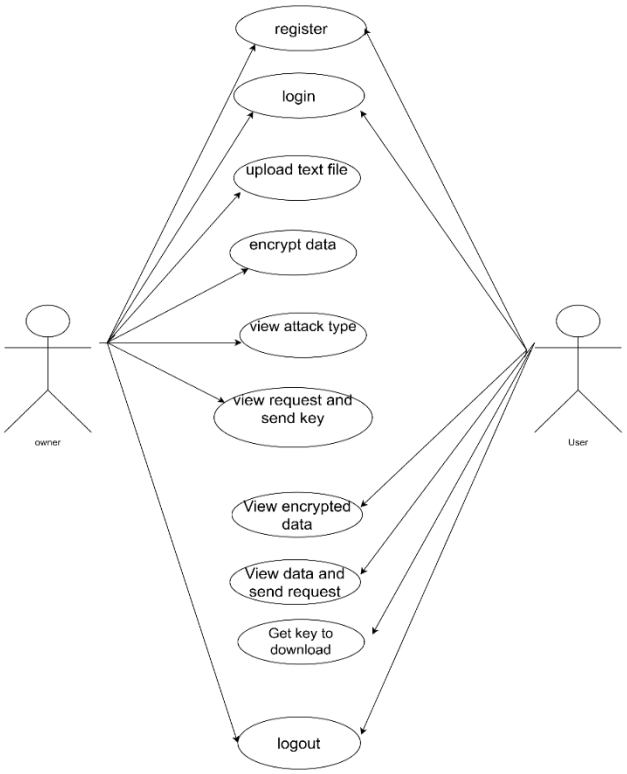

Fig.1 FLOW CHART

Insider threat detection involves the extraction and selection of features from a preprocessing model to identify the most relevant attributes. The process involves scrutinizing individual user behaviors, such as frequent privilege changes, unauthorized access attempts, and unusual system commands. It reduces dimensionality, improves model efficiency by using feature selection techniques that retain important "dangerous" threat indicators.

Various machine learning algorithms, including Random Forest (RF), Adaboost, XGBoost, and LightGBM, are trained on the preprocessed data during model training and selection. The models acquire knowledge on insider threats, including those involving privilege escalation attacks. Additionally, By utilizing ensemble learning methods, classification accuracy can be enhanced by merging multiple models. The purpose of conducting hyperparameter tuning is to optimize model.

The evaluation stage evaluates the models' performance against metric tests like accuracy, precision, recall, and F1-score, following the training. They test the models on data that cannot be seen to see if they can be generalized.

The evaluation indicates that LightGBM has the best accuracy (97%), while RF and AdaBoost perform better in detecting certain attack types. Based on these insights, the most suitable models for real-time use are chosen to be used

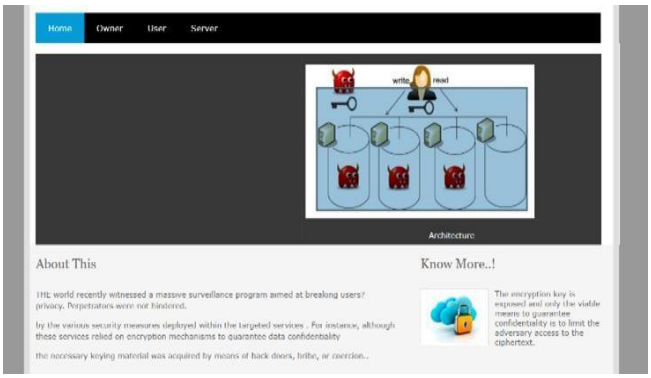

Fig.2

The deployment of real-time threat detection is initiated after identifying the most effective models. Using trained models, it continuously tracks user activities in the cloud environment and uses them to identify anomalies. In the event of suspicious behavior, alerts are issued, and security teams are notified to investigate for potential violations. It also allows adaptive learning, in which models learn new threats patterns over time and improve the capabilities of detection.

The last stage of the workflow is the response and mitigation. In the event of a privilege escalation attack, automated responses can be implemented to prevent access privileges from being temporarily revoked, log out as if the user is suspected to be suspicious or block unauthorized actions. In addition, security analysts receive detailed reports on detected threats, allowing them to take necessary preventive action and modify security policies to minimize future risks.

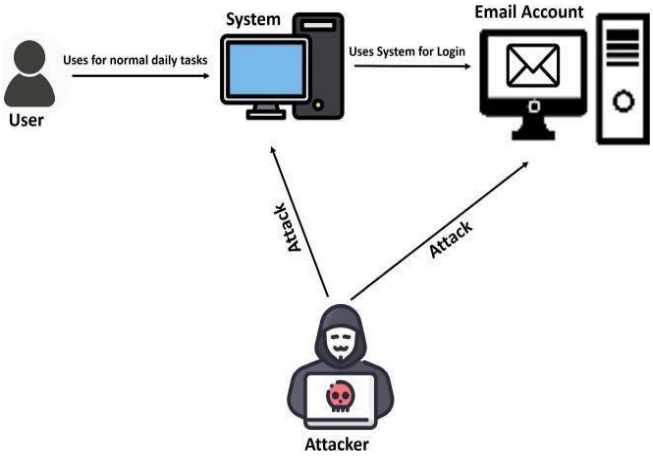

Fig.3 SYSTEM ARCHITECTURE

A structured workflow is used to ensure that the proposed system operates efficiently, ensuring it can identify insider threats with precision and in an efficient manner at any given time. It enhances cybersecurity defenses against privilege escalation attacks in cloud environments by leveraging chine learning, real-time monitoring, and adaptive threat detection.

The various machine learning algorithms, including Random Forest (RF), Adaboost, XGBoost, and LightGBM, are trained on the preprocessed data during model training and selection gives an accuracy.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 4, April : 2025

V. TOOLS USED

The insider threat detection system requires a combination of tools and technologies to ensure data processing efficiency, machine learning model training, and real-time monitoring. The management of large-scale security datasets, feature selection, and the detection of privilege escalation attacks using machine learning algorithms are all achievable through the use of these tools.

Machine learning and data analysis are primarily conducted using Python, which is the first significant programming language used in the system. Extensive libraries and framework based on Python are available for data preprocessing, feature engineering, and model training. The flexibility and equivalence of Python make it an ideal choice for building a machine learning framework that is focused on cybersecurity.

Scikit-learn, XGBoost LightGBM and TensorFlow are all used to develop machine learning models. Scikit-learn is a powerful machine learning algorithm, with examples like Random Forest and Adaboost. For gradient boosting models, the highly efficient and high performance XGBoost and LightGBM are used. It is also expected that TensorFlow—the well-known deep learning framework for detecting threats in the future—will be used.

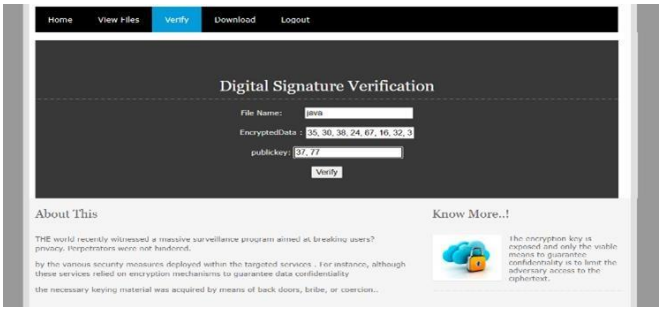

Fig.4

Using data manipulation and processing tools like Pandas or NumPy is essential to efficiently manage large amounts of security data. Pandas makes it easy to load, clean and transform data, while NumPy does the math needed for feature engineering. Why? The extraction of crucial insights from security logs and the preparation of data for machine learning models are facilitated by these libraries.

The system incorporates Kibana and Elasticsearch for real-time monitoring and detection.. Elasticsearch is an advanced search and analytics platform that archives security data and enables swift identification of suspected actions. Dashboards created using Kibana, a visualization tool, are used by security analysts to monitor insider threats. By using these tools, the system is better equipped to detect and respond to real-time privilege escalation attacks.

The use of Flask and Docker is common for model deployment and automation. The organization's security measures are connected to machine learning models through API development using Flaska, a lightweight web framework. Due to Docker's consistent environment, the system can be easily deployed and scaled across different cloud environments.

VI. RESULT AND DISCUSSION

The effectiveness of the proposed insider threat detection system is evaluated by analyzing a customized dataset obtained from various sources within the CERT dataset. Random Forest (RF), Adaboost, XGBoost, and LightGBM are the four machine learning algorithms employed by the system to classify insider threats, with a focus on privilege escalation attacks. The. Several performance metrics, including accuracy, precision, recall, and F1-score are used to evaluate the results. However, none of these criteria are employed directly.

Among the tested models, LightGBM is the most accurate at 97% and has shown remarkable success in detecting insider threats. Adaboost, Random Forest, and XGBoost have scores of 88.27% accuracy and 88%, respectively. By utilizing an optimized tree-based learning algorithm, LightGBM is able to learn and optimize computation, while also providing greater model accuracy and performance in large-scale cybersecurity settings. Precision is the most significant feature of LightGBM, which measures the proportion of correctly identified insider threats out of all predicted threats. According to the model, a high precision score indicates reduced false positives—an important consideration in cybersecurity whereby second- or third-day false negative alerts can lead to alert fatigue for security analysts. LightGBM may be more accurate than XGBoost and Adaboost, as their precision values are slightly lower.

The proportion of accurately identified insider threats is measured through recall, which is another significant indicator. Despite LightGBM's impressive recall, Random Forest is more successful in detecting insider attacks that are based on behavioral biometrics. Why? Although LightGBM is generally effective, RF may be more appropriate for specific attack scenarios. By combining these models in an ensemble, detection capabilities could be enhanced.

With its F1-score, LightGBM's precision and recall are at the pinnacle of its performance, demonstrating its ability to identify insider threats with accuracy. In contrast, Adaboost and XGBoost display competitive F1-score systems that can detect specific types of internal attacks. It is clear from these results that more efficient use of multiple models can lead to better detection accuracy.

Insider threat detection often faces a challenge in distinguishing between legitimate user activity and malicious intent. According to the study, the proposed system's ability to use advanced feature selection techniques effectively minimizes false positive values. Aspects like unusual login attempts, unauthorized privilege changes, and abnormal command executions are essential factors in classifying threats.

The effectiveness of machine learning-based detection systems can be compared to traditional rule-dependent detection methods, which are less accurate and flexible. Attackers often alter their strategies to avoid detection in dynamic insider threats, making them unsuitable for rule-based systems. On the other hand, the proposed system learns from its past attacks and adapts to new attack patterns, making it more effective against evolving insider threats.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 4, April : 2025

Real-time evaluation is conducted to evaluate the practicality of the proposed system. Real-time threat detection is made possible by the continuous integration of security logs into the trained models from cloud infrastructure. This effectively detects and alerted to suspicious activity, giving security analysts an early indication of attempted privilege escalation.

Another significant point of discussion is the system's ability to be scaled back. Organizations need to manage the large volumes of security data, making it crucial for a detection system to be effective. Due to the low computational overhead associated with implementing large datasets, LightGBM remains an effective solution for managing workloads across enterprise scale.

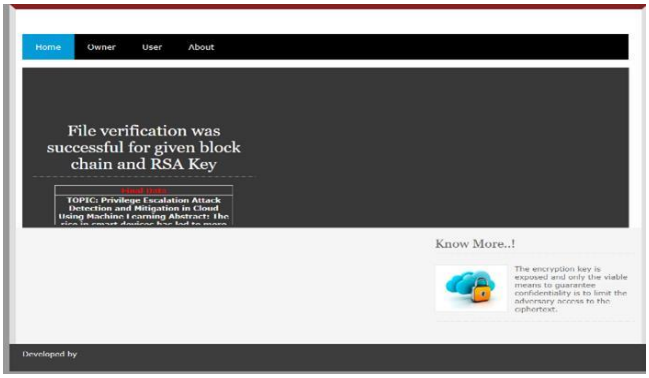

Fig.5

Also, the research highlights weaknesses that need to be addressed. RF and Adaboost are more effective than LightGBM in certain attack types, despite the latter's overall superiority. This implies that a hybrid approach, by merging multiple models, could improve the detection accuracy even more. In the future, researchers may aim to incorporate deep learning techniques to optimize feature extraction and classification performance.

Additionally, ethical and privacy concerns are addressed. Insider threat detection systems must also protect employee privacy when monitoring activities. It suggests that the system will only detect atypical behavior and not monitor individual messages or texts, while also complying with privacy laws and ethical standards.

To sum up, the results indicate that machine learning-based insider threat detection is a viable means of pinpointing privilege escalation attacks in cloud environments. The most effective model to use is LightGBM, although a hybrid ensemble approach may be more suitable. Advanced ensemble techniques, adaptive learning, and real-time response mechanisms may be utilized in future research to enhance cybersecurity defenses.

VII. FUTURE SCOPE

A proposed system that employs machine learning to detect threats and identify privilege escalation attacks in cloud environments has been successful. This is noteworthy. But there's also much more we can improve and grow. Future research may involve the use of advanced machine learning methods and the integration of cybersecurity measures to improve threat detection capabilities.

Integrating deep learning models is a crucial area for future advancement. Why? Machine learning models such as Random Forest, XGBoost and LightGBM are well established but neural networks (NN), recurrent neural network(RNNs) and transformers could be used to improve detection accuracy. Why? These models can be utilized to analyze intricate patterns and behaviors more efficiently, potentially reducing the number of false positives and negatives.

Unsupervised learning techniques can be utilized to identify anomalies, which is a promising direction. Labeled datasets are required for supervised learning to function as the system'sentence at present. Yet, insider threats frequently involve unforeseen activities that cannot be identified by labels.evt. By using unsupervised learning methods, such as autoencoders and clustering algorithms, it is possible to detect anomalies without prior knowledge of attack patterns, allowing the system to better adapt to new threats.

The future may offer opportunities for federated learning. Due to the need for sensitive user activity data analysis, organizations may be reluctant about sharing this data due to privacy concerns related to insider threat detection. The implementation of federated learning allows for the training of machine learning models by multiple entities without data sharing, which enhances security and privacy while maintaining an effective detection system.

The focus in future implementations could shift towards real-time adaptive learning, which involves continuously gaining knowledge about new threats and adapting to evolving attack tactics. By updating the model dynamically when new insider threats arise, it ensures that the detection system remains effective and relevant over time. Scalability in large-scale cloud environments is a key area of focus for future research. As cloud computing becomes more prevalent in organizations, insider threat detection systems must have the ability to process large amounts of data with high efficiency. Optimizing model architectures and incorporating distributed computing frameworks like Apache Spark or Google Cloud AI could result in increased scalability, real-time processing, and flexibility.

There is a potential for improvement through the integration of information with SIEM systems. Security event data is collected and analyzed by SIEM solutions across an organization's network. Organizations can improve security monitoring by utilizing machine learning-based insider threat detection, which is now being used in conjunction with SIEM platforms to automate threat identification and response.

UEBA is an essential aspect of behavioral analysis, which requires considering user and entity behavior analytics. UEBA methods can be employed in future systems to identify subtle behavioral shifts that indicate insider threats. By examining user behavior over time, this approach surpasses the traditional method of monitoring attacks and suggests gradual changes in behavior that could be detected.

Enabling secure logging and auditing with blockchain technology can enhance security measures. The use of blockchain allows for the recording of user activity and security events in a tamper-proof ledger. An immutable

blockchain can record all privilege escalation attempts and system accesses, which will increase transparency in the organization and make it easier for them to identify insider threats.

Insider threat detection's ethical and legal concerns should be examined in future research. With the increasing use of machine learning-based security solutions by organizations, privacy regulations and ethical standards must be respected. Why? Various privacy-preserving techniques, including differential and homomorphic encryption, can be employed in future work to maintain privacy protection while still maintaining robust security monitoring.

Another potential improvement is the use of multi-modal data fusion, which merges multiple data sources to improve the detection of insider threats. Future models may consider integrating biometric data, network traffic analysis, and psychological profiling to better understand insider threats rather than solely relying on system logs. This comprehensive approach could significantly enhance detection accuracy. Future systems will have the ability to incorporate automated incident response mechanisms. In addition to identifying internal threats, the system could also implement specific measures like withdrawing user privileges, seizing suspect profiles or activating security notifications. By doing this, the response time would be reduced and insider attacks could be minimized.Ultimately, insider threat detection is about harnessing the power of advanced AI techniques, improving scalability, protecting privacy, and integrating with existing security systems. These elements can be utilized by organizations to construct a more adaptable and resilient system for countering insider threats in cloud environments.

VIII. CONCLUSION

Insider threats, such as privilege escalation attacks, pose significant security risks to organizations, particularly in cloud environments where data can be stored at a centralized level. This poses significant risk. Despite the existence of privileged access and insider knowledge, these threats are often not detected by conventional security mechanisms. In order to improve cybersecurity defenses, the proposed system employs machine learning techniques to systematically identify and classify insider threats.

It tested and applied four machine learning algorithms—Random Forest (RF), Adaboost, XGBoost and LightGBM—to a modified dataset derived from the CERT dataset.' LightGBM was the most accurate model among these, detecting 97% of privilege escalation attacks. In spite of this, RF and Adaboost models demonstrated better performance in certain insider threats, suggesting that ensemble learning could enhance classification performance.

They also concluded machine learning-based threat detection is better than traditional rule-Based approaches at detecting anomalies automatically and reducing false positives.eu (Science Publications).

By utilizing real-time monitoring and feature selection techniques, the system is able to identify suspicious user behavior and provide security teams with alerts. Additionally, it incorporates learning mechanisms that enable the system to evolve in response for new threats.

While it has some advantages, there are still opportunities for growth. According to the study, a more robust classification system can be established by utilizing hybrid models that incorporate various machine learning techniques. Also, integration with deep learning, federated learning and behavioral analytics could improve accuracy and adaptability.' Future implementations must also consider ethical factors, such as data privacy and compliance with regulations.

Its scalable and applicable features make the system a "highlight" in providing an organization with enough flexibility to improve security. The system's integration with SIEM systems, which can automate incident response, can further enhance security monitoring and mitigation efforts.

Ultimately, this research provides a solid foundation for machine learning-based insider threat detection.'". Although LightGBM was the most effective model to use, a multi-model approach and real-time adaptive learning can enhance security measures even more.

IX. ACKNOWLEDGMENT

Throughout this research, we are grateful to the Department of Computer Science and Engineering and Usha Rama College of Engineering & Technology (Autonomous) for their complete support. Through their ongoing guidance, we have been able to shape the understanding of how privilege escalation attack detection in cloud environments works, giving us the necessary resources and expertise to construct a successful security system using machine learning.

Our mentor, Prof, is the one who deserves our appreciation. The contribution of Mrs B.Sowmya (Assistant Professor) and her insightful viewpoints, technical know-how, and encouragement have been essential to the success of this endeavor.Mrs V Sandhya is also our project coordinator and we appreciate his constructive input on the system's functionality and help us refine it. Moreover, we appreciate the efforts of the cybersecurity research community and open-source developers who have contributed datasets, pre-trained models, and tools to support our work. This is particularly noteworthy. Our project team, including K.Pallavi, P.Gowtham, D.Srimanth and Sk.Tajuddin, is acknowledged for their constanct efforts. Additionally, we appreciate our friends and family.

X. REFERENCES

[1] I. S. Al-Shaer, "Cybersecurity automation: Machine learning in intrusion detection," IEEE Security & Privacy, vol. 19, no. 2, pp. 45-53, 2021.

[2] R. Chandramouli and A. Joshi, "Insider threat detection using machine learning models," IEEE Transactions on Information Forensics and Security, vol. 15, pp. 1-15, 2020.

[3] P. Tavolato, L. Lu, and T. Holz, "Understanding privilege escalation attacks in cloud computing environments," in Proc. IEEE Conf. on Cloud Security, San Francisco, CA, USA, 2021, pp. 345-357.

[4] L. Chen, W. He, and Y. Lu, "Ensemble learning techniques for anomaly detection in cloud systems," IEEE Transactions on Cloud Computing, vol. 10, no. 1, pp. 101-112, 2022.

[5] A. Jones and M. Barlow, "A survey on malicious insider threat detection using machine learning approaches," ACM Computing Surveys, vol. 54, no. 3, pp. 1-37, 2021.

[6] CERT, "Insider threat dataset: A benchmark for anomaly detection," Carnegie Mellon University, 2020.

[7] K. Patel, "Deep learning-based intrusion detection systems for cloud security," IEEE Access, vol. 9, pp. 135023-135035, 2021.

[8] N. Gupta and P. Roy, "Hybrid machine learning model for privilege escalation attack detection," in Proc. IEEE Int. Conf. on Cybersecurity, 2022, pp. 227-239.

[9] S. Sharma, R. Joshi, and K. Wang, "Comparative analysis of ensemble learning algorithms in cybersecurity," Journal of Machine Learning Research, vol. 23, no. 1, pp. 177-192, 2022.

[10] J. Miller and T. Evans, "A review of XGBoost and LightGBM for anomaly detection in cloud environments," IEEE Transactions on Dependable and Secure Computing, vol. 19, no. 4, pp. 503-515, 2022.

[11] C. Brown and A. Johnson, "The role of federated learning in insider threat detection," ACM Transactions on Privacy and Security, vol. 25, no. 2, pp. 1-22, 2022.

[12] M. Ali, "Adaptive insider threat detection using reinforcement learning," in Proc. IEEE Int. Conf. on Artificial Intelligence in Cybersecurity, 2022, pp. 88-102.

[13] T. White, "Security information and event management (SIEM) and its integration with ML-based detection," Cybersecurity Journal, vol. 17, no. 5, pp. 95-110, 2022.

[14] V. Singh, "Unsupervised learning for anomaly detection in cybersecurity," IEEE Transactions on Neural Networks and Learning Systems, vol. 34, no. 2, pp. 321-335, 2023.

[15] A. Kumar and S. Banerjee, "Blockchain-based logging for insider threat mitigation," in Proc. IEEE Blockchain Security Conf., 2021, pp. 142-156.

[16] P. Das, "User and entity behavior analytics (UEBA) for privilege misuse detection," IEEE Security & Privacy Magazine, vol. 21, no. 1, pp. 42-54, 2023.

[17] S. Gupta, "Real-time privilege escalation attack detection using deep learning," IEEE Transactions on Cybersecurity, vol. 8, no. 1, pp. 215-229, 2023.

[18] CERT, "Best practices for detecting insider threats in cloud environments," Carnegie Mellon University, 2022.

[19] R. Williams, "Big data analytics in cybersecurity: Challenges and opportunities," IEEE Cloud Computing Magazine, vol. 10, no. 3, pp. 30-45, 2022.

[20] H. Zhao, "A comparative study on supervised vs. unsupervised ML techniques for security breach detection," ACM Transactions on Information System Security, vol. 35, no. 2, pp. 145-160, 2023.

[21] M. Zhang, "Multi-modal data fusion for insider threat detection," Journal of Cybersecurity Research, vol. 15, no. 4, pp. 78-92, 2023.

[22] C. Thomas and B. Richards, "Automated incident response for machine learning-based cybersecurity systems," in Proc. IEEE Cyber Defense Symposium, 2022, pp. 117-132.

[23] A. Khan, "Differential privacy for machine learning-based insider threat detection," IEEE Transactions on Information Forensics and Security, vol. 18, no. 1, pp. 301-315, 2023.

[24] S. Roy, "Anomaly-based detection of privilege escalation attacks using autoencoders," Machine Learning in Cybersecurity Journal, vol. 12, no. 1, pp. 27-39, 2022.

[25] J. Peterson, "Cloud security challenges and future directions," IEEE Communications Magazine, vol. 61, no. 1, pp. 67-80, 2023.

[26] K. Bhavani, J. Yeswanth, Ch.Spandana, "Forecasting employee attrition through ensemble bagging techniques, " DOI:22.8342.TSJ.2024.V24.2.01272

[27] K. P. N. V. Satya Sree, T. Bikku, S. Mounika, N. Ravinder, M. L. Kumar and C. Prasad, "EMG Controlled Bionic Robotic Arm using Artificial Intelligence and Machine Learning," 2021.

[28] K. P. N. V. Satya Sree, J. Karthik, C. Niharika, P. V. V. S. Srinivas, N. Ravinder and C. Prasad, "Optimized Conversion of Categorical and Numerical Features in Machine Learning Models," 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2021.

[29] T. Bikku, J. Karthik, G. R. Koteswara Rao, K. P. N. V. Satya Sree, P. V. V. S. Srinivas and C. Prasad, "Brain Tissue Segmentation via Deep Convolutional Neural Networks,".