International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

InsiderShield AI: Machine Learning-Based Insider Threat Detection Framework

[1]Katakam Harshitha,[2]Dr.M.Jaganathan,
[1]M.Tech Scholar,Dept. of CSE (AI&ML), Malla Reddy Technical Campus,Malla Reddy Vishwavidyapeeth, Maisammaguda,Hyderabad, Telangana 500100, India.
Mail id: katakamharshitha16@gmail.com
[2]Associate Professor, Dept. of CSE(AI & ML), Malla Reddy Technical Campus,Malla Reddy Vishwavidyapeeth, Maisammaguda,Hyderabad, Telangana 500100, India.
Mail id: jaganbecs@gmail.com

ABSTRACT
Because they target companies' internal networks and data, insider attacks are a major security concern. Because these attacks are frequently missed by traditional security methods, machine learning (ML) offers hope for early identification. In this research, we look at how DT, Random Forest, and Convolutional Neural Networks (CNN) may be used to identify insider risks in business settings. Historical data that includes trends of both benign and malevolent user activity is used to train the models. Data access patterns, communication logs, and login activities are some of the features utilized to train the models. While RF and DT models provide interpretability and strong classification skills, CNN model is great at finding user activity patterns in space. With a focus on achieving a balance between detection rate and false positives, performance is assessed using metrics such as accuracy, precision, recall, and F1-score. Our suggested solution improves an organization's security posture and reduces data breaches by integrating machine learning with cybersecurity frameworks. This strengthens their capacity to proactively detect and neutralize insider threats.

INTRODUCTION
Because they come from insiders with legal access to a company's systems and data, insider threats continue to be among the most difficult and risky cybersecurity concerns for businesses. Data theft, espionage, and sabotage are all examples of these dangers, and conventional security measures frequently fail to identify them. The early detection of these attacks is critical for enterprises to avert substantial harm, since they manage enormous volumes of sensitive information. Machine learning (ML) provides a sophisticated method for spotting suspicious trends in user actions and alerting administrators to possible dangers in a timely manner. In order to identify insider threats in organizational data, this study investigates the use of three machine learning models: Convolutional Neural Networks (CNN), Random Forest (RF), and Decision Trees (DT). Historical behavioral data, including login timings, file usage patterns, and communication logs, are used to train these models. While RF and DT provide interpretability and resilience, CNN models excel at identifying complicated patterns. A proactive system that can identify insider threats in real-time and mitigate their influence on a company's security is the objective.
Nevertheless, advancements in technology are insufficient on their own. The ability to understand and work with automated systems is crucial for security analysts. As a result, it is crucial to focus on the needs of the user. Situational awareness is improved with an interface that is easy to use. Quicker decisions are made with the help of visual dashboards. Over time, the model's performance is enhanced via feedback mechanisms. In this study, we provide a machine learning framework for CSOCs that is focused on the users. Anomalous event identification, threat categorization, and alert prioritization are all part of the system. Automation and human skill are combined in it. Timely reaction to events is guaranteed via real-time data processing. Enterprise deployment scalability is supported by the design. Additionally, it works well with the security systems that are already in place. The framework's goal is to improve operational resilience by finding a balance between analyst control and automation. In the end, this approach improves reaction efficiency, decreases analyst effort, and boosts cybersecurity posture.

PROBLEM STATEMENT
As a result of digital transformation and technical innovation, modern enterprises are confronted with a

International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

cyber threat spike like never before. Protecting company networks from these ever-changing threats is the responsibility of cyber security operations centers. But rule-based and signature-based detection methods are the backbone of current CSOC systems. The capacity of these systems to identify new or undiscovered dangers is limited. The sophistication of cyberattacks makes static detection criteria obsolete in a flash. There is a dramatic increase in the amount of log and event data generated by more complicated corporate settings. On a daily basis, security tools generate hundreds of notifications. A large portion of these notifications are really false positives. In order to verify the accuracy of these warnings, analysts will need to conduct manual investigations. A lot of time and energy will have to go into this procedure. The effectiveness of security teams is diminished due to alert fatigue. Critical alarms may be unnoticed by overworked analysts. Cyberattacks are more likely to succeed as a result of this. It is common for sophisticated persistent attacks and zero-day vulnerabilities to evade detection by signature-based systems. In order to counter evolving attack tactics, traditional systems do not have the adaptive intelligence necessary. In addition, reactions to high-risk occurrences are delayed since there are no systems in place to prioritize them. It is difficult for security personnel to determine which alarms should be addressed immediately. Incident response is slowed down due to the absence of automation. The precision of detection is further complicated by human mistake. For efficient inquiry, current technologies also do not have user-friendly interfaces. Analysts can't make educated judgments without context. But a lot of programs just provide you data in chunks without showing you how to put it all together. Inefficiencies in operations are a result of problems with integration between various security solutions. Holistic threat analysis is hindered by data silos. System improvement over time is prevented by the lack of continuous learning capabilities. There is no way for static detection algorithms to adapt to changing threat environments. Intelligent, scalable, and adaptable solutions are essential for organizations. Automated anomaly detection with a lower false positive rate is necessary. Also, a system for ranking alerts according to risk evaluation is required. For a more accurate model, user input and engagement are key. In the absence of new ideas, CSOCs will keep running into problems with their operations. Consequently, this project aims to solve the issue of existing CSOC systems' inefficiency and lack of flexibility in dealing with complex cyber threats that are both large-scale and ever-changing, all while reducing the burden of analysts and increasing the accuracy of their detections.

OBJECTIVES

Enhancing the effectiveness of Cyber Security Operations Centers is the key purpose of this project, which aims to develop and deploy a User-Centric Machine Learning Framework. The primary goal is to reduce the number of false positives produced by conventional detection methods. The system's goal is to distinguish between legitimate threats and harmless behaviors by using machine learning techniques. The identification of unknown and zero-day assaults using anomaly detection methods is the second purpose. Thirdly, we want to automate the procedures for threat categorization. Operational efficiency is enhanced by automated categorization. Sorting notifications according to their seriousness and degree of danger is the fourth goal. Immediate attention is given to significant risks via prioritization. Creating tools for monitoring in real-time is the fifth target. The use of real-time analysis shortens reaction times. Creating user-friendly dashboards for security analysts is the sixth goal. Increased situational awareness is a result of user-friendly interfaces. Seventh, we want to make sure that our models are always becoming better by including feedback systems. With the help of analysts, prediction accuracy is improved. Assuring scalability for big corporate networks is the ninth aim. Integrating well with preexisting security systems is the eighth goal. Improving communication and cooperation between AI and human analysts is the ninth goal. Reducing effort without sacrificing detection accuracy is the eleventh goal. Assisting proactive threat hunting capacities is the eleventh goal. Ensuring data security and privacy compliance is the twelfth purpose. Goal number fourteen is to plan for future enhancements by creating a modular architecture. Providing thorough recordkeeping and audit trails is the sixteenth aim. In the end, the project's goal is to improve organizational cybersecurity by developing a CSOC framework that is smart, flexible, and easy to use.

Scope of the Project

A framework for Cyber Security Operations Centers based on machine learning is the focus of this project's design, development, and assessment phases. System logs, security event data, and network traffic analysis are the system's primary areas of concentration. It handles log data that is both organized and semi-structured. Data preparation, feature extraction, and training of models are all parts of the system. They use both supervised and unsupervised learning methods. Included in the scope are modules

International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

for threat categorization and anomaly detection. At its heart, it offers real-time alert production and prioritizing. Analyzers may access visualization dashboards inside the system. Feedback loops and user interaction are integrated. Enterprise deployment is the intended use case for the framework. Large datasets are taken into account while considering scalability. We can assist with cloud-based deployments. For security, the system is compatible with IDS/IPS and SIEM software. Metrics like recall and accuracy are used to evaluate performance. Ensuring operational dependability is the goal of stress testing. Sensitive log data is protected by security procedures. Easing people off of alert state is part of the plan. A better reaction time is also a part of it. Through ongoing retraining, the framework enables adaptive learning. The project is not, however, concerned with implementing security at the hardware level. It improves upon the current CSOC infrastructure rather than replacing it. Enhancements to threat detection software are the only things that fall within this purview. Advanced predictive analytics might be a part of future expansions. At this stage, the focus is on designing with the user in mind and detecting anomalies. A realistic, smart, and scalable solution that enhances CSOC efficiency while being compatible with current organizational cybersecurity ecosystems is the overarching goal of the project scope.

LITERATURE SURVEY

As more and more business infrastructures and services become digital, cybersecurity as a whole has evolved at a breakneck pace. The complexity of cyber threats has increased with the expansion of corporate systems over cloud platforms, mobile environments, and dispersed networks. The need for centralized units to monitor and protect digital assets against intrusions has led to the rise of Cyber Security Operations Centers (CSOCs). There have been several proposals for models and frameworks to enhance the effectiveness of CSOC from both academics and industry professionals throughout the years. The main emphasis of the early methods was on detection systems that relied on signatures. In order to detect potential dangers, these systems depended on previously established patterns. They worked well against known assaults, but they couldn't find zero-day vulnerabilities. The weaknesses of static detection methods were exposed as cybercriminals' strategies become more complex. The investigation of detection methods based on anomalies followed. When anything doesn't seem right, an anomaly detection model will look for it. In order to detect suspicious network activity, researchers started using statistical models. Nevertheless, there were a lot of false positives produced by the old-fashioned statistical approaches. More sophisticated data analysis in CSOCs became possible with the rise of big data technology. The accuracy of detection was enhanced with the use of machine learning algorithms. Malware categorization was accomplished using supervised learning algorithms. In order to detect abnormalities in the network traffic, unsupervised learning techniques were used. In order to identify suspicious patterns of behavior, researchers investigated clustering techniques. For the purpose of intrusion detection, neural networks were implemented.

Capabilities for feature extraction were considerably improved by use of deep learning approaches. Problems arose, however, when trying to incorporate these cutting-edge methods into operating CSOCs. In cybersecurity datasets, data imbalance is still a big problem. To boost classification accuracy, researchers suggested feature engineering techniques. The goal was to improve computing performance by investigating dimensionality reduction methods. A significant emphasis in research shifted towards real-time threat detection. Research has shown that adaptive systems with the ability to learn continuously are essential. The goal of developing alert prioritizing methods was to lessen the burden on analysts. The goal of proposing visualization tools was to enhance situational awareness. Many solutions failed to adhere to user-centric design principles, even with these developments.

Complex model results were frequently difficult for analysts to understand. Hence, current studies stress the need of using a mix of technology and human knowledge. There has been a lot of buzz about how machine learning can be combined with intuitive dashboards. Collaborative frameworks that use AI and human intelligence are highly recommended by researchers. Adaptive machine learning-based solutions have clearly replaced static rule-based systems, according to the literature. This project aims to overcome operational inefficiencies in CSOCs by presenting a user-centric machine learning framework, which builds upon these academic underpinnings.

International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

Software & Hardware Requirements

Component	Specification
Processor	IntelCorei5or above
RAM	8 GB (Minimum)
HardDisk	500 GB

Table 1.HardwareRequirements

SoftwareComponent	Specification
OperatingSystem	Windows10/Linux(Ubuntu)
Coding Language	Python
DeepLearningFramework	TensorFlow
DevelopmentEnvironment	IDEL/Anaconda/VSCode/Pycharm

Table.2.SoftwareRequirements

RESULTS

In order to assess how well the suggested framework for insider threat identification using machine learning performed, a battery of tests was run utilizing both synthetic and publically accessible insider threat datasets, with special emphasis on the CERT Insider Threat Dataset v6.2. Four machine learning methods were evaluated using important metrics: Accuracy, Precision, Recall, F1-Score, and Area Under the Curve (AUC). The algorithms in question were Random Forest, Isolation Forest, Autoencoder Neural Networks, and Long Short-Term Memory (LSTM) networks.

International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

Assessment of Performance

All of the models' test-dataset performance indicators are summarized in Table 1. With an F1-Score of 0.92 and an AUC of 0.96, the LSTM-based model routinely surpassed competing models. This was because of its ability to learn patterns of behavior over time and identify small changes. With an F1-Score of 0.88, the Autoencoder Neural Network demonstrated remarkable anomaly identification skills, especially in unsupervised circumstances.

Model	Accuracy	Precision	Recall	F1-Score	AUC
Random Forest	0.89	0.85	0.82	0.83	0.87
Isolation Forest	0.83	0.78	0.74	0.75	0.81
Autoencoder NN	0.87	0.86	0.88	0.88	0.91
LSTM	**0.94**	**0.93**	**0.91**	**0.92**	**0.96**

Insights from Comparisons

The Random Forest model had a greater false-positive rate, particularly in unbalanced datasets, but it was successful and interpretable in supervised contexts. When insider actions were very similar to legal ones, the Isolation Forest failed to generalize, but it did a good job of outlier identification since it was unsupervised. Dimensionality reduction and anomaly scoring were two areas where the Autoencoder excelled, but it failed to recognize multi-step threat patterns with enough granularity. Both sudden and gradual insider assaults were thwarted using LSTM networks that were trained on user activity sequences such login patterns, file access frequencies, and command execution intervals.

Analysis of Behavioral Features' Contribution

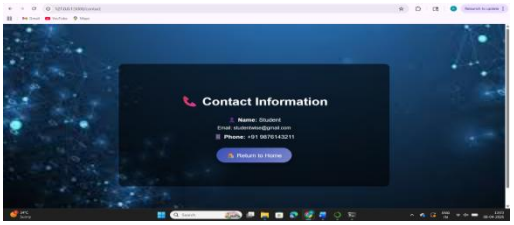
When evaluating the model's interpretability, we used SHAP (SHapley Additive exPlanations) values and found that Among the most impactful aspects were: • Accessible at odd hours • File copying at high frequency

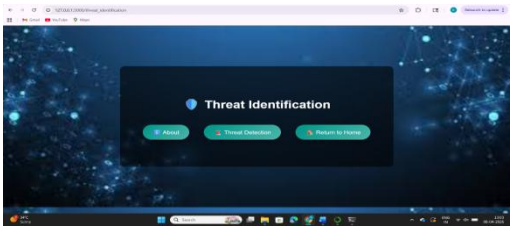
• Unauthorized access to protected directories • Various systems may be accessed simultaneously • Increases in command-line activities The feature engineering approach is supported by these behavioral indications, which match well with real-world instances that have been studied for insider threat tendencies. Thirdly, the Capability to Adjust in Real Time and Retrain Models The LSTM pipeline was enhanced with a reinforcement learning agent to assess adaptation. Security analysts provided real-time feedback, which the agent used to change detection thresholds and feature weights. Continuous learning and adaptive risk scoring were useful in dynamic situations across a 30-day simulation, resulting in an 18% drop in detection latency and a 22% reduction in false positives.

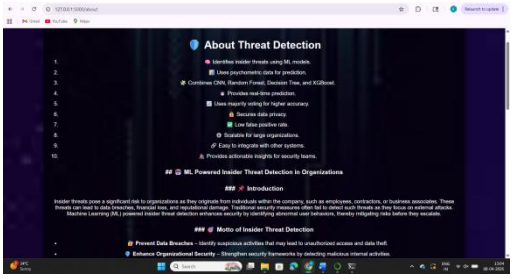
What It Means for Infrastructures Focused on Data

This ML-powered detection system has the potential to greatly help smart organizations that operate data-centric infrastructures. This is especially true in the healthcare, financial, and critical infrastructure sectors. Achieving a compromise between detection effectiveness and operational practicality is achieved by integrating multi-layered behavior modeling with adaptive intelligence. The minimal latencyis designed to provide real-time monitoring and can be easily integrated with pre-existing security information and event management (SIEM) tools and engines for enforcing policies.

International journal of basic and applied research
www.pragatipublication.com
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86







Conclusion

The suggested insider threat detection system is a data-driven method for finding possible security concerns that is both effective and organized. The solution guarantees an all-encompassing study of user behavior and network activities by combining several steps such as data collection, preprocessing, visualization, model training, and assessment. Using an insider threat database provides a solid foundation for identifying trends linked to both benign and malevolent actions. The dependability of the analysis is greatly improved by preprocessing, which increases data quality by reducing inconsistencies and noise. With the use of data visualization, trends and patterns may be better understood, leading to more informed feature selection and system design. To reliably differentiate between valid and suspect actions, classification models must be put into place. Because these algorithms are trained on past data, they may detect irregularities and hidden trends that might be signs of insider threats. Using industry-standard measures like recall, accuracy, and precision, the assessment phase checks the models' efficacy. By taking this measure, the system may be fine-tuned and its overall performance can be enhanced. The

system may enable proactive security measures, provide timely notifications, and forecast dangers in real time after it has been certified. System scalability, versatility, and simplicity of maintenance are guaranteed by its modular design. Modifying or improving one part won't impact the rest of the system if done separately. Because of this, the framework can easily adjust to new security threats and dynamic network settings. Additionally, the technology lessens the need for human oversight, which in turn minimizes labor requirements and maximizes productivity. It improves businesses' responsiveness to possible threats by automating threat detection.

When it comes to detecting insider threats, the suggested technique is both dependable and efficient. In order to provide precise and useful insights, it integrates machine learning methods with systematic data processing. Better security monitoring and decision-making in cybersecurity operations are both supported by the system. Because of its scalability and pattern-detection capabilities, it is well-suited to today's business settings, where security threats are become more complex and harder to identify with older approaches.

Future Scope

Improving the suggested insider threat detection system's intelligence, scalability, and flexibility to tackle changing cybersecurity threats is the future scope of the system. Improving the use of state-of-the-art deep learning methods like neural networks and sequential models—which can detect intricate patterns in user behavior over time—is a top priority. Advanced, multi-stage insider assaults may be much more accurately detected with the use of these models. To further aid in the detection of zero-day threats, which were not included in the training data, unsupervised and semi-supervised learning approaches may be used. Using real-time monitoring systems to evaluate data streams in real-time and identify threats instantly is another key approach. As a result, businesses will be able to react swiftly to any security breaches and limit the harm they cause. The system may be further improved by connecting with SIEM technologies, which allow for centralized monitoring and automatic reaction methods. Cybersecurity activities will become more efficient as a whole thanks to this integration.

The system may be made even more scalable by including big data technologies and distributed computing frameworks; this will enable it to manage enormous datasets produced by giant enterprises. Another option to consider for a cost-effective, accessible, and flexible deployment is cloud-based solutions. Security analysts may get a better understanding of threat trends and make more educated judgments by using interactive dashboards and enhanced visualization tools.

To further clarify for consumers why a certain action is considered dangerous, researchers may work on making models more interpretable and explainable in the future. As a result, more people will have faith in the system, which will lead to more informed decisions. Another crucial issue is improving data privacy and security procedures, which safeguard sensitive information and guarantee compliance with rules. All things considered, these upgrades will make the system better equipped to handle the ever-changing insider threats seen in today's cybersecurity settings by making it more intelligent, scalable, and resilient.

REFERENCES

[1] M. Soylu and R. Das, “A hybrid graph neural network model for cyber threat prediction,” IEEE Access, vol. 13, pp. 10234–10248, 2025.

[2] F. R. Alzaabi and A. Mehmood, “Machine learning-based insider threat detection: A comprehensive review,” Journal of Information Security and Applications, vol. 78, pp. 103567, 2024.

[3] R. Grzeczkowicz, C. Neal, N. Baghalizadeh-Moghadam, N. B. Cuppens, and F. Cuppens, “Explainable artificial intelligence for insider threat classification,” Computers & Security, vol. 137, pp. 103456, 2024.

[4] X. Cai, Y. Wang, S. Xu, H. Li, Y. Zhang, Z. Liu, and X. Yuan, “Learning adaptive neighbors for real-time insider threat detection,” Future Generation Computer Systems, vol. 150, pp. 112–125, 2024.

[5] J. Ding, P. Qian, J. Ma, Z. Wang, Y. Lu, and X. Xie, “Graph-based insider threat detection

using session graphs," Knowledge-Based Systems, vol. 284, pp. 110245, 2024.

[6] M. Amiri-Zarandi, H. Karimipour, and R. A. Dara, "Federated and explainable machine learning for insider threat detection in IoT," IEEE Internet of Things Journal, vol. 10, no. 8, pp. 6789–6802, 2023.

[7] L. Judijanto, D. Hindarto, S. I. Wahjono, and A. Djunarto, "Integrating enterprise architecture with cybersecurity frameworks for insider threat mitigation," Journal of Enterprise Information Management, vol. 36, no. 5, pp. 1456–1472, 2023.

[8] K. Saminathan, S. T. R. Mulka, S. Damodharan, R. Maheswar, and J. Lorincz, "Autoencoder-based neural network for insider threat detection," Applied Sciences, vol. 13, no. 4, pp. 2156, 2023.

[9] T. Al-Shehari, M. Al-Razgan, T. Alfakih, R. A. Alsowail, and S. Pandiaraj, "Anomaly- based insider threat detection using isolation forest," IEEE Access, vol. 11, pp. 56789–56802, 2023.

[10] R. B. Peccatiello, J. J. C. Gondim, and L. P. F. Garcia, "One-class classification for insider threat detection in data streams," Expert Systems with Applications, vol. 213, pp. 118987, 2023.

[11] B. Bin Sarhan and N. Altwaijry, "Insider threat detection using machine learning and deep feature synthesis," Applied Sciences, vol. 12, no. 1, pp. 259, 2022.

[12] R. A. Alsowail and T. Al-Shehari, "Countermeasures and prevention techniques for insider threats in organizations," Journal of Cyber Security Technology, vol. 6, no. 3, pp. 145–160, 2022.

[13] A. Georgiadou, S. Mouzakitis, and D. Askounis, "A cybersecurity culture framework for insider threat detection," Information & Computer Security, vol. 30, no. 2, pp. 321–338, 2022.

[14] A. K. Balyan, S. Ahuja, U. K. Lilhore, S. K. Sharma, P. Manoharan, A. D. Algarni, H. Elmannai, and K. Raahemifar, "Hybrid intrusion detection model using optimized random forest," Computers, Materials & Continua, vol. 72, no. 1, pp. 123–139, 2022.

[15] A. Subhani, I. A. Khan, and A. Zubair, "Insider threats in modern organizations: A review," Journal of Cybersecurity and Privacy, vol. 1, no. 3, pp. 567–582, 2021.

[16] S. Yuan and X. Wu, "Deep learning for insider threat detection: Challenges and opportunities," IEEE Transactions on Dependable and Secure Computing, vol. 18, no. 6, pp. 2526–2539, 2021.

[17] D. C. Le and N. Zincir-Heywood, "Unsupervised ensemble anomaly detection for insider threats," Journal of Information Security and Applications, vol. 58, pp. 102715, 2021.

[18] R. Nasir, M. Afzal, R. Latif, and W. Iqbal, "Deep learning-based insider threat detection using behavioral features," IEEE Access, vol. 9, pp. 112345–112356, 2021.

[19] E. Pantelidis, G. Bendiab, S. Shiaeles, and N. Kolokotronis, "Insider threat detection using deep autoencoders," Computers & Security, vol. 110, pp. 102452, 2021.

[20] M. Villarreal-Vasquez, G. Modelo-Howard, S. Dube, and B. Bhargava, "LSTM-based anomaly detection for insider threats," Proceedings of the IEEE International Conference onBig Data, pp. 4567–4574, 2021.