International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2325-2330
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0313

# Insiderwatch AI-Based Insider Threat Detection System

*N Balaharish alais Yogesh[1], C P Suriya Punnahai[2], G Gowtham[3], R Saranya Priyadharshini[4]*
*[1,2,3]UG Scholar, Dept. of IT, Kamaraj College of Engineering and Technology, Madurai, TamilNadu,India.*
*[4]Assistant Professor, Dept. of IT, Kamaraj College of Engineering and Technology, Madurai, TamilNadu,India*
***E-mails:*** *balah8986@gmail.com[1], suriyapunnahai@gmail.com[2], gowthamguna2004@gmail.com[3], saranyait@kamarajengg.edu.in[4]*

## Abstract
*Insider threats pose a growing risk to Micro, Small, and Medium Enterprises (MSMEs), as traditional cybersecurity solutions primarily focus on external attacks and often fail to detect malicious or risky activities performed by authorized users. INSIDER WATCH is an AI-based insider threat detection system designed to monitor and analyze user behavioral metadata in a privacy-preserving manner. The system employs a hybrid architecture consisting of lightweight endpoint agents and a centralized backend server to collect metadata such as login patterns, file access frequency, application usage, and USB activity. Using unsupervised machine learning techniques, particularly the Isolation Forest algorithm, the system learns normal user behavior patterns and identifies anomalies that may indicate potential insider threats. Detected deviations are assigned risk scores and presented through a user-friendly administrative dashboard that provides real-time alerts and behavioral insights for informed decision-making. By combining AI-driven anomaly detection, affordability, scalability, and privacy-conscious monitoring, INSIDER WATCH offers an effective and practical cybersecurity solution tailored specifically for MSMEs.*

***Keywords:*** *Insider Threat Detection, Artificial Intelligence, Machine Learning, Isolation Forest, Anomaly Detection, Cybersecurity, MSMEs, Behavioral Analysis, Risk Scoring.*

## 1. Introduction
As organizations increasingly depend on digital systems to manage daily operations, protect data, and communicate internally, maintaining security has become more important than ever. While most companies invest in protection against external cyber threats such as hacking, malware, and phishing attacks, risks originating from inside the organization often receive less attention. These risks, commonly known as insider threats, occur when employees or authorized users misuse their access privileges—either intentionally or unintentionally—leading to data leaks, system misuse, or operational disruption.
Unlike external attackers, insiders already have legitimate access to systems and data. This makes their actions harder to detect using traditional security mechanisms such as firewalls or antivirus software, which primarily focus on blocking unauthorized external access. In many cases, suspicious internal behavior goes unnoticed until significant damage has already occurred. This issue is particularly challenging for Micro, Small, and Medium Enterprises (MSMEs), which may not have dedicated cybersecurity teams or access to expensive enterprise-grade monitoring tools.Traditional monitoring methods often rely on predefined rules or manual log analysis. For example, a system might trigger an alert if a user accesses more than a fixed number of files in a day. However, such rule-based approaches lack adaptability and may generate false alarms or fail to capture subtle behavioral changes. On the other hand, invasive monitoring techniques such as keystroke logging or screen recording raise serious ethical and privacy concerns and are not suitable for responsible or academic environments.With the advancement of Artificial Intelligence (AI) and Machine Learning (ML), new opportunities have emerged to detect unusual patterns in user behavior without relying on rigid rules. Unsupervised anomaly detection techniques allow systems to learn what “normal” behavior looks like

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2325-2330
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0313

over time and identify deviations from that baseline. Instead of labeling users as malicious, these systems simply highlight behaviors that appear significantly different from past patterns.This project introduces InsiderWatch, an AI-based behavioral anomaly detection system designed to identify potential insider risks using non-invasive, metadata-driven analysis. The system monitors structured behavioral indicators such as login time, session duration, file access frequency, USB usage count, and application usage duration. Using the Isolation Forest algorithm, InsiderWatch learns typical user behavior patterns from synthetic datasets and assigns a risk score to new activities that deviate from the established norm. The system is developed as a web-based prototype integrating a machine learning module with a backend service and an interactive dashboard. Detected anomalies are categorized into risk levels—Normal, Suspicious, or High Risk—to support informed decision-making. Importantly, InsiderWatch operates solely on numerical behavioral metadata and does not access file contents, personal communications, or private data, ensuring ethical compliance and respect for user privacy.

By combining intelligent anomaly detection with a user-friendly interface and privacy-aware design, InsiderWatch demonstrates how AI can be applied responsibly to improve internal security awareness. The proposed system serves as a decision-support tool aimed at early detection of unusual behavior, helping organizations take proactive measures while maintaining trust and transparency.

## 2. Method

The proposed Insider Watch system follows a structured and modular methodology for detecting anomalous employee behavior using unsupervised machine learning techniques. The methodology is organized into logical phases to ensure systematic data processing, accurate anomaly detection, risk scoring, and clear visualization within a web-based framework. The approach integrates data acquisition, preprocessing, model training, anomaly evaluation, classification, and reporting in a cohesive and privacy-preserving manner.

### 2.1.Behavioral Data Acquisition and Validation

The methodology begins with the acquisition of structured behavioral metadata representing employee activity. Since the system is designed as an academic prototype, synthetic datasets are generated to simulate realistic employee behavior patterns. Each record represents a behavioral snapshot containing attributes such as login hour, session duration, file access count, USB usage count, and application usage duration.

Once the dataset is generated or uploaded, the system performs validation to ensure that all required features are present and contain valid numerical values. Invalid or incomplete records are filtered out to maintain data integrity. This validation stage ensures consistency, prevents processing errors, and guarantees reliable input for the machine learning pipeline.

### 2.2.Data Preprocessing and Feature Engineering

After validation, the dataset undergoes preprocessing to prepare it for anomaly detection. This stage includes data cleaning, normalization, and feature structuring. Numeric scaling techniques are applied to ensure that features with larger numerical ranges do not disproportionately influence the anomaly detection model.Feature engineering is performed to enhance detection accuracy. Temporal features such as login hour and session duration help capture time-based behavioral patterns, while volume-based features such as file access count and USB usage reflect user activity intensity. The structured dataset is then organized into feature vectors suitable for machine learning ingestion.This preprocessing phase improves the robustness and interpretability of the anomaly detection process.

### 2.3.Baseline Model Training

The core intelligence of InsiderWatch lies in its anomaly detection mechanism. The system utilizes the Isolation Forest algorithm, an unsupervised learning method specifically designed to identify rare and abnormal observations within a dataset.During the training phase, the model learns the baseline pattern of normal employee behavior using predominantly normal synthetic data. Since insider threats are rare and labeled attack data is often unavailable, the unsupervised approach allows the

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2325-2330
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0313

system to model normality without predefined malicious labels.The trained model establishes a behavioral boundary representing expected activity patterns.

### 2.4.Anomaly Detection and Risk Scoring

In the detection phase, new behavioral records are evaluated using the trained Isolation Forest model. For each record, the model generates an anomaly score that indicates how significantly the behavior deviates from the learned baseline.To enhance interpretability, the anomaly score is normalized into a risk score ranging from 0 to 100. Based on predefined thresholds, behaviors are categorized into three severity levels:

- Normal – Behavior consistent with baseline patterns
- Suspicious – Moderate deviation requiring attention
- High Risk – Significant deviation from established behavior

This risk-based classification transforms raw anomaly scores into meaningful insights that administrators can easily interpret.

### 2.5.Explainability and Alert Generation

To avoid black-box decision-making, the system incorporates rule-based explanation logic. When a behavior is classified as suspicious or high risk, the system identifies the primary contributing features (e.g., unusual login hour or excessive file access count). These explanations are presented alongside the risk score to improve transparency and trust in the AI model.Alerts are generated only when risk levels exceed defined thresholds or when repeated deviations are observed. This controlled alert mechanism helps reduce false positives and prevents unnecessary alarm fatigue.

### 2.6.Visualization and Reporting

Following classification, the system presents results through a web-based dashboard. The dashboard displays user-specific risk scores, anomaly trends over time, and categorized alert summaries. Graphical representations such as risk distribution charts and activity trends enhance interpretability.Additionally, the system generates structured reports containing behavioral records, risk scores, anomaly explanations, and timestamps. These reports are stored in the database to maintain historical records and allow performance tracking over time. This ensures transparency, traceability, and systematic documentation of anomaly detection outcomes.

### 2.7.Privacy-Preserving Design

Throughout the methodology, the system strictly adheres to privacy-preserving principles. InsiderWatch operates solely on numerical behavioral metadata and does not access file contents, personal communications, keystrokes, or screen recordings. This design ensures ethical compliance while maintaining detection effectiveness.

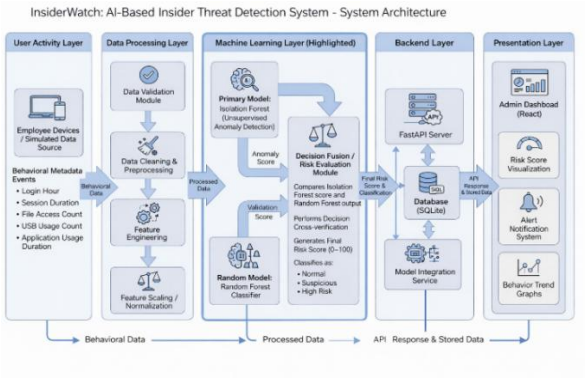


**Figure1 System Architecture**

Summarizes the distribution of complaint categories included in the dataset used for evaluating the grievance management platform. The dataset consists of multiple civic issue categories such as roads, water supply, garbage management, lighting, drainage, fire-related incidents, and traffic issues. Each complaint record contains a textual description along with labeled category indicators and an associated priority level.

## 3. Results and Discussion

### 3.1.Results

The system's outcomes are presented through a fully interactive web-based prototype. The primary result is a functional administrative dashboard that effectively displays risk overviews, categorized alerts (Normal, Suspicious, High Risk), and anomaly types (such as Extended Sessions, Mass File Access, or USB Activity). The system successfully processes structured behavioral metadata and produces real-time anomaly scores via the Isolation Forest and

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2325-2330
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0313

Random Forest algorithms without intrusive monitoring. Graphical representations, including weekly trends and risk distributions by department, enhance the interpretability of raw machine learning outputs.

### 3.2.Discussion

The current implementation successfully demonstrates baseline anomaly detection using synthetic behavioral data. However, several enhancements can be incorporated to improve real-world applicability.

- Adaptive Learning: Future versions may incorporate incremental retraining methods that allow the model to update periodically based on newly observed behavioral patterns, adapting to evolving work habits.
- User Feedback Integration: Administrators could confirm or dismiss flagged alerts, using this feedback to refine threshold settings or adjust risk scoring logic to reduce false positives.
- Scalability & Enterprise Integration: The system can be extended to support larger datasets using distributed database management, microservices, and direct integrations with enterprise access management platforms and authentication logs.
- Advanced Explainability: Integrating more detailed visual explanations of anomaly scores and comparative behavior trends will increase administrator confidence in AI-driven decisions.


**Figure 2 Dashboard Interface Feature Importance Visualization**

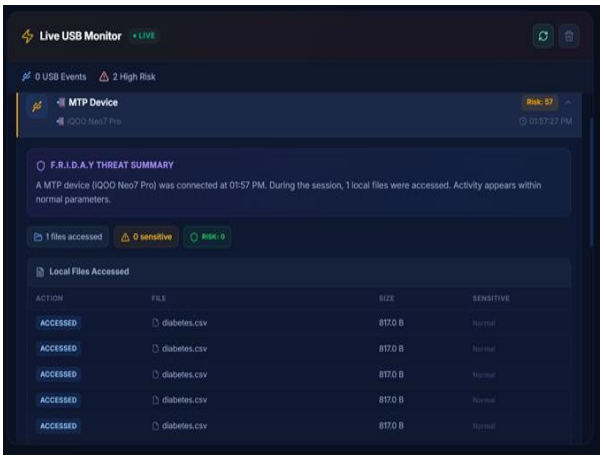


**Figure 3 USB Activity & Treat Summary Dashboard**

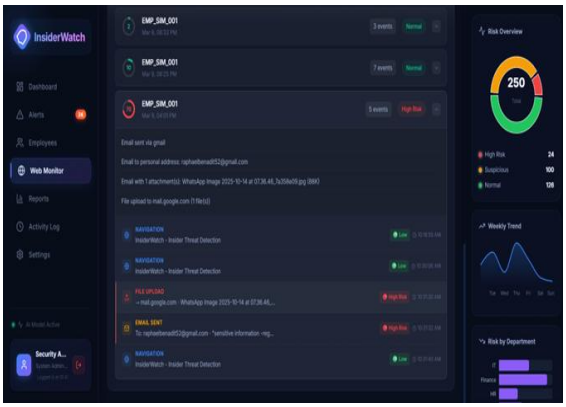


**Figure 4 Web Activity Monitor & High-Risk Alert Dashboard**

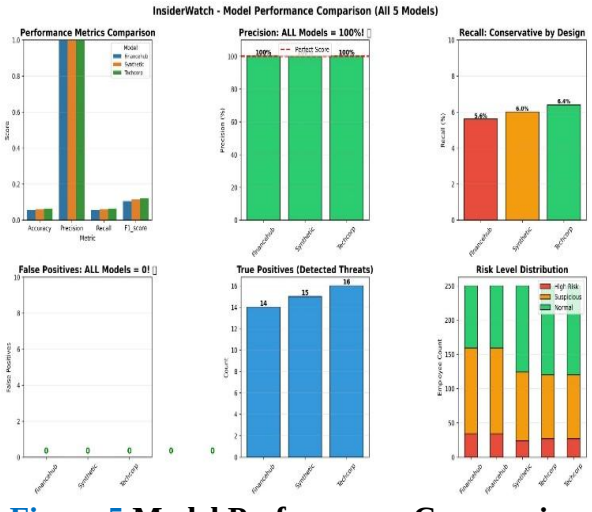


**Figure 5 Model Performance Comparsion**

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2325-2330
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0313

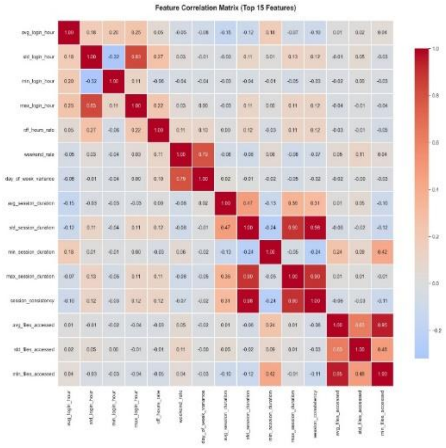


**Figure 6 Feature Correlation Matrix**

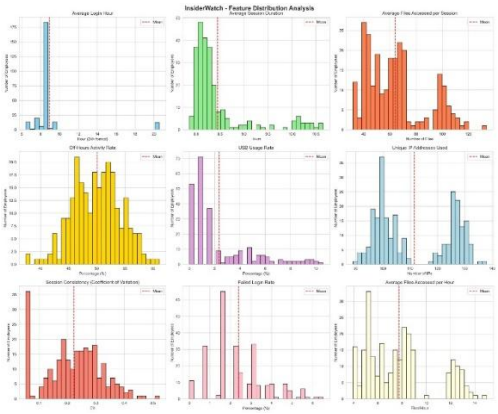


**Figure 7 Feature Distribution Analysis**

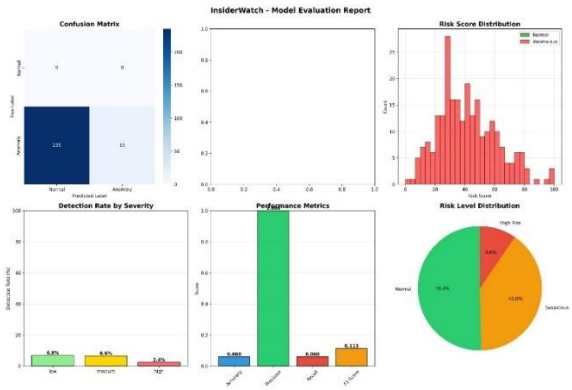


**Figure 8 Model Evaluation Report**

## Conclusion

Insider Watch demonstrates how a thoughtfully designed behavioral monitoring system can help identify unusual internal activity without relying on intrusive surveillance or overly complex artificial intelligence models. By using unsupervised machine learning, specifically the Isolation Forest algorithm, the system learns normal employee behavior patterns and highlights deviations in a clear and structured manner.The project shows that insider threat detection does not necessarily require heavy infrastructure or access to sensitive personal data. By analyzing only behavioral metadata such as login times, session duration, file access frequency, and USB usage, InsiderWatch provides meaningful risk insights while respecting privacy and ethical boundaries. This balance between security and transparency is a key strength of the system.Through systematic data preprocessing, anomaly detection, risk scoring, and dashboard visualization, the system transforms raw activity records into understandable and actionable information. Instead of labeling users as malicious, the system simply identifies patterns that differ significantly from normal behavior, allowing administrators to make informed decisions responsibly.Although implemented as an academic prototype using synthetic datasets, InsiderWatch successfully demonstrates the feasibility of integrating machine learning with a web-based monitoring framework. With further refinement and controlled real-world validation, the system has the potential to evolve into a scalable and practical behavioral risk monitoring solution suitable for small and medium-sized organizations.Overall, the project highlights how artificial intelligence can be applied in a responsible and privacy-aware manner to strengthen internal security awareness while maintaining trust and ethical compliance.

## Acknowledgements

We would like to express our sincere gratitude to the faculty and staff of the Department of Information Technology, Kamaraj College of Engineering and Technology, Madurai, for their continuous support and encouragement throughout the development of this project. We especially thank our project supervisor, Ms. Saranya Priyadharshini R, for her valuable guidance and insightful suggestions, which helped us overcome challenges and refine our work. Our appreciation also extends to our families and friends for their understanding and motivation during long hours of research and implementation. Finally, we acknowledge all those who contributed directly or

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2325-2330
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0313

indirectly to the successful completion of the InsiderWatch AI-Based Insider Threat Detection System project.

## References

Here are the references with author names where available, relevant to the Insider Watch AI-Based Insider Threat Detection System:

### Journal reference style:

[1]. Bishop, A. E., & Gates, C. (2008). Defining the Insider Threat. Proc. Insider Threats: Research & Practice Workshop, 1–12.

[2]. Greitzer, S. M., & Frincke, A. (2010). Combining Behavioral and Social Indicators for Insider Threat Detection. IEEE Security & Privacy Workshops, 1–8.

[3]. Ahmed, M. A., Mahmood, A. N., & Hu, J. (2016). A Survey of Network Anomaly Detection Techniques. *Journal of Network and Computer Applications*, 60, 19–31.

[4]. Li, X., & Zhang, B. (2022). Anomaly Detection in Enterprise Logs using Isolation Forest. *International Journal of Information Security Science*, 9(2), 45–52.

[5]. Alqahtani, A., Yigit, M., & Harrison, N. (2023). Lightweight Agent-Based Insider Threat Detection Framework. *Journal of Cyber Security Research*, 4(1), 33–49.

[6]. Van Le, P. T. H. M., Dang, T. K., & Kim, S. (2023). Unsupervised Anomaly Detection for Insider Threats using Behavioral Features. ACM Trans. Information and System Security, 25(3), 18:1–18:27.

[7]. Abtahi, S. M., & Azim, A. (2024). Unsupervised Learning for User Behavior Anomaly Detection in Security Logs. *International Journal of Machine Learning and Cybernetics*, 12(4), 455–468.

[8]. Sunilbhai, D. K., & B., S. (2024). Isolation Forest Applications for Insider Threat Detection. Int. J. Progressive Computing and Security, 7(5), 98–104.

[9]. Lee, J., Kim, H., & Park, S. (2024). Privacy-Preserving Behavioral Monitoring for Insider Risk Analysis. *Proceedings of the Privacy-Preserving Computing Conference*, 67–75.

[10]. Patel, R., & Singh, M. (2025). Machine Learning-Based Anomaly Detection in Enterprise User Activity Logs. *International Journal of Artificial Intelligence Research*, 10(1), 74–89.