

---

# Journal of Informatics and Web Engineering

Vol. 4 No. 2 (June 2025)

eISSN: 2821-370X

---

## Anomaly Detection in Network Traffic for Insider Threat Identification: A Comparative Study of Unsupervised and Supervised Machine Learning Approaches

Sellappan Palaniappan<sup>1\*</sup>, Rajasvaran Logeswaran<sup>2</sup>, Shapla Khanam<sup>3</sup>

<sup>1</sup>Corporate Office, HELP University, No. 15, Jalan Sri Semantan 1, Off Jalan Semantan, Bukit Damansara 50490 Kuala Lumpur, Malaysia

<sup>2,3</sup>Faculty of Computing and Digital Technology, HELP University, Persiaran Cakerawala, Subang Bestari, 40150 Shah Alam, Selangor, Malaysia

\*corresponding author: (sellappan.p@help.edu.my; ORCID: 0009-0009-1168-2864)

**Abstract** - Insider threats pose a significant and growing risk to organizational cybersecurity, with recent studies indicating a 47% increase in insider incidents from 2018 to 2022. This paper presents a comparative analysis of unsupervised and supervised machine learning approaches for detecting potential insider threats through network traffic anomaly identification. We develop and evaluate an Isolation Forest (unsupervised) and a Random Forest (supervised) model, training them on a simulated dataset representing six months of network logs from a mid-sized company. Our study introduces a unique feature set combining traditional network metrics with temporal and behavioural indicators, enhancing the models' detection capabilities. Results show that the Random Forest classifier outperforms the Isolation Forest, with F1-scores of 0.6425 and 0.4624, respectively. However, the unsupervised approach shows promise in scenarios lacking labelled data. Key findings reveal that increased connection frequency and data transfer volume are critical indicators of potential threats, with temporal patterns also playing a significant role. This study provides valuable insights into the strengths and limitations of each approach, offering practical implications for real-world digital forensics investigations. We contribute to the field by proposing a hybrid approach that leverages the strengths of both methods, potentially improving the accuracy and adaptability of insider threat detection systems. These findings pave the way for more robust, context-aware cybersecurity measures in the digital age.

**Keywords**— Insider Threat Detection, Network Security, Machine Learning, Anomaly Detection, Digital Forensics

Received: 29 August 2024; Accepted: 28 December 2024; Published: 16 June 2025

This is an open access article under the [CC BY-NC-ND 4.0](#) license.



---

### 1. INTRODUCTION

Cybersecurity threats come from both within and without organizations. Unlike external attacks, insider threats originate from within the organization's network, and this makes threat detection more challenging because they have legal access to corporate resources. According to the 2023 Insider Threat Report by Cybersecurity Insiders, 74% of organizations feel vulnerable to insider threats, with 39% reporting an increase in insider incidents over the past 12

months [1]. This alarming trend underscores the growing importance of developing effective insider threat detection mechanisms.

Conventional or traditional rule-based detection systems do not always detect insider threats as insiders are typically authorized users who have access to their organization's resources. Such systems are weak in differentiating between normal and malicious behavioural patterns when insiders use their own legitimate access privileges to conduct malicious activities. Recent studies in [2-5] highlight the effectiveness of using advanced machine learning techniques in cybersecurity applications, including for intrusion detection and malware classification. The adaptability and pattern recognition capabilities of machine learning algorithms allow them to identify the subtle anomalies in insider threats.

The process of detecting and mitigating insider threats is costly, in terms of financial resources as well as the potential damage to an organization's reputation. In a recent study by the Ponemon Institute, it was reported that the average cost of insider threats rose by 34% from 2020 to 2022, reaching \$15.38 million per incident [6]. This highlights the sophistication of insider threats, and the vital need for the development of effective detection and prevention methods. Organizations are now becoming increasingly aware of the importance of investing in advanced threat detection systems to safeguard their digital assets, to maintain their competitive edge in an increasingly digital landscape.

This study sets out to evaluate the effectiveness of supervised and unsupervised machine learning algorithms in detecting insider threats, specifically by detecting anomalies in network traffic within an organization. With the results obtained, this paper aims to shed light on the strengths and limitations of each approach and provide valuable insights to researchers and practitioners in this field.

## 2. RESEARCH QUESTIONS AND OBJECTIVES

Our study aims to evaluate the application of machine learning techniques for insider threat detection, focusing on the comparison of supervised and unsupervised approaches. This study will address the following research questions:

- RQ1 - What features identify network traffic anomalies that are indicative of insider threats?
- RQ2 - How to generate an appropriate dataset for model evaluation?
- RQ3 - How do supervised and unsupervised machine learning algorithms compare in terms of their performance in detecting insider threats within an organization's network?
- RQ4 - How can these machine learning insights be integrated into real-world cybersecurity strategies?

The following are the research objectives of this work, catered to answer the research questions stated:

- RO1 - To develop a simulated dataset representing network traffic with insider threat scenarios.
- RO2 - To identify the most influential features for insider threat detection.
- RO3 - To evaluate the practical ability of Isolation Forest and Random Forest models in anomaly detection.
- RO4 - To propose insights to integrate cybersecurity strategies against insider threats into existing real-world frameworks.

## 3. LITERATURE REVIEW

In recent literature, research in insider threat detection has increasingly focused on machine learning approaches due to the growing complexity of cyber threats [6], [7]. Tuor et al. proposed an unsupervised deep learning model for insider threat detection [8], while Chattopadhyay et al. developed a scenario-based system emphasizing contextual information [2]. More recent advances in deep neural networks with attention mechanisms have shown promising results in this field [9]. It has been demonstrated in the literature that ensemble methods like Random Forest often outperform single-model approaches in supervised learning scenarios [3]. The effectiveness of Isolation Forest in identifying rare anomalies in network intrusion detection was showcased in [10] and is a pertinent point to be considered. Additionally, the importance of feature engineering in improving model accuracy, as emphasized in [7], should also be considered when developing any machine learning solution. Recent studies in [4], [5], [11] highlight the effectiveness of using advanced machine learning techniques in cybersecurity applications, including for intrusion detection and malware classification.

Despite the advancements in recent times, there is still a gap in terms of a lack of comparisons between unsupervised and supervised approaches for insider threat detection, particularly in network traffic analysis. Most existing research focuses on either supervised or unsupervised methods, with limited comparative performance evaluation in practical scenarios. Such an evaluation would be useful in strategizing the development of future cybersecurity solutions.

## 4. METHODOLOGY

In this work, simulation would be used to evaluate the performance of unsupervised and supervised machine learning models in detecting insider threats via analysis of network traffic patterns. From the literature, it was determined that the best unsupervised model for this work was the Isolation Forest algorithm, as described in [12]. Recent simulation studies have further confirmed the effectiveness of unsupervised machine learning algorithms for anomaly detection [13], [14]. Random Forest classifier, as outlined in [15], was selected for the supervised model.

### 4.1 Dataset Simulation

To evaluate the models, a dataset representing 6 months of network traffic logs from a mid-sized company was simulated to reflect the real-world scenario. The dataset was generated to contain 99% normal traffic patterns and 1% simulated insider threat activities, reflecting the rare nature of insider threats in real-world scenarios. This imbalance was deliberately introduced to test the models' ability to detect rare events. The dataset includes the following features:

- Employee ID: a unique identifier for each employee
- Timestamp: date and time of the network activity
- Source and Destination IP: simulated IP addresses for network connections
- Protocol: types of network protocols used (HTTP, HTTPS, FTP, SSH)
- Bytes transferred: volume of data transferred in each connection
- Time of day and Day of week: temporal features derived from the timestamp
- Department: simulated organizational departments (HR, IT, Finance, Sales, Marketing)
- Access level: simulated security clearance levels (1-5, with 5 being highest)
- Duration of connection: time spent on each network activity
- Number of connections: frequency of network activities per employee

### 4.2 Data Preprocessing

The data was pre-processed as follows:

- Feature engineering: We created time-based features from the timestamp data, including `time_of_day` and `day_of_week`. These temporal features are crucial for capturing patterns in user behaviour that may indicate insider threats.
- Normalization: Numerical features were standardized using `StandardScaler` from `scikit-learn`. This step ensures that all features are on the same scale, preventing features with larger magnitudes from dominating the model training process.
- Encoding: Categorical variables, such as 'department', were one-hot encoded using `OneHotEncoder`. This transforms categorical data into a format suitable for machine learning algorithms.
- Balancing: To address the class imbalance in our dataset, we applied for the Synthetic Minority Over-sampling Technique (SMOTE) for supervised model training. This technique creates synthetic examples of the minority class (insider threats) to balance the dataset.

### 4.3 Model Implementation

We implemented two models for anomaly detection, namely, an Isolation Forest (unsupervised model) and a Random Forest classifier (supervised model). Pseudocodes 1 and 2 are the brief description of these models respectively.

## Pseudocode 1. Isolation Forest

---

```

#Isolation Forest (unsupervised model):
function IsolationForest(X, n_trees, sample_size):
    forest = []
    for i = 1 to n_trees:
        X_sample = RandomSample(X, sample_size)
        tree = BuildIsolationTree(X_sample)
        forest.append(tree)
    return forest

function BuildIsolationTree(X):
    if X.size == 1 or TreeHeight == HeightLimit:
        return ExternalNode(X.size)
    else:
        q = RandomSelectFeature(X)
        p = RandomSplitPoint(X, q)
        X_left, X_right = SplitData(X, q, p)
        return InternalNode(
            BuildIsolationTree(X_left),
            BuildIsolationTree(X_right),
            SplitAttribute = q,
            SplitValue = p
        )

function AnomalyScore(x, forest):
    path_length = 0
    for tree in forest:
        path_length += PathLength(x, tree)
    return 2^(-path_length / len(forest) / c(len(X)))

function Predict(X, forest, threshold):
    anomalies = []
    for x in X:
        score = AnomalyScore(x, forest)
        if score > threshold:
            anomalies.append(x)
    return anomalies

```

---

The Isolation Forest algorithm works by recursively partitioning the data space, isolating anomalies in regions with fewer data points. Anomalies typically result in shorter path lengths in the trees, leading to higher anomaly scores. We used the `IsolationForest` class from `scikit-learn` with the parameters `contamination` set to 0.01, reflecting our simulated 1% insider threat rate, and `random_state` set to 42 for reproducibility.

The Random Forest is an ensemble learning method that constructs multiple decision trees and combines their outputs. Each tree in the forest votes on the classification of a data point, with the majority vote determining the final prediction. We used the `RandomForestClassifier` from `scikit-learn` with the parameters `n_estimators` set to 100 trees and `random_state` set to 42 for reproducibility.

Pseudocodes 1 and 2 provide a high-level overview of how each algorithm operates. In the implementation, `scikit-learn` library shown in pseudocode 3 provide optimized version of these algorithms:

## Pseudocode 2. Random Forest

---

```

# Random Forest Classifier (Supervised):
function RandomForest(X, y, n_trees, max_depth):
    forest = []
    for i = 1 to n_trees:
        X_sample, y_sample = BootstrapSample(X, y)
        tree = BuildDecisionTree(X_sample, y_sample, max_depth)
        forest.append(tree)
    return forest

```

---

---

```

function BuildDecisionTree(X, y, max_depth):
    if Pure(y) or max_depth == 0:
        return LeafNode(MajorityClass(y))
    else:
        best_feature, best_split = FindBestSplit(X, y)
        X_left, X_right, y_left, y_right = SplitData(X, y, best_feature, best_split)
        left_subtree = BuildDecisionTree(X_left, y_left, max_depth - 1)
        right_subtree = BuildDecisionTree(X_right, y_right, max_depth - 1)
        return InternalNode(left_subtree, right_subtree, best_feature, best_split)

function Predict(X, forest):
    predictions = []
    for x in X:
        votes = []
        for tree in forest:
            votes.append(PredictTree(x, tree))
        predictions.append(MajorityVote(votes))
    return predictions

function PredictTree(x, tree):
    if IsLeaf(tree):
        return tree.prediction
    else:
        if x[tree.split_feature] <= tree.split_value:
            return PredictTree(x, tree.left_child)
        else:
            return PredictTree(x, tree.right_child)

```

---

### Pseudocode 3. Using scikit-learn Library

---

```

from sklearn.ensemble import IsolationForest, RandomForestClassifier

# Isolation Forest
isolation_forest = IsolationForest(contamination=0.01, random_state=42)
isolation_forest.fit(X_train)

# Random Forest
random_forest = RandomForestClassifier(n_estimators=100, random_state=42)
random_forest.fit(X_train, y_train)

```

---

## 5. RESULTS AND DISCUSSIONS

### 5.1 Model Performance

Figure 1 shows graphically the performance metrics for Isolation Forest and Random Forest models. The performance metrics for the two models are shown in Table 1.

Table 1. Model Performance

Isolation Forest	Random Forest
Precision: 0.4674	Precision: 0.6765
Recall: 0.4574	Recall: 0.6117
F1-Score: 0.4624	F1-Score: 0.6425

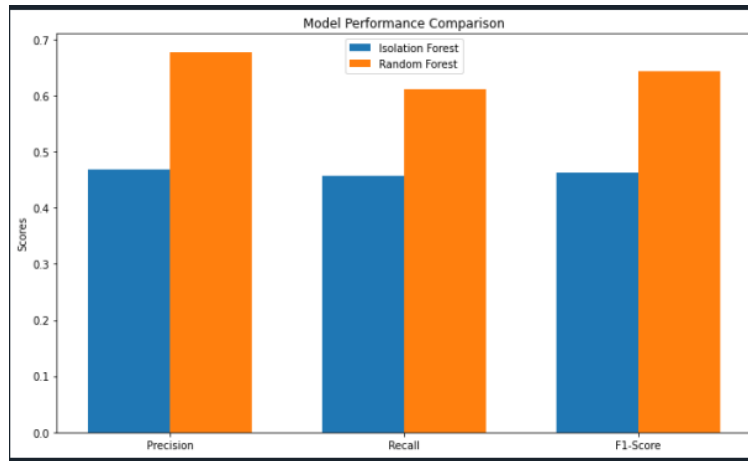


Figure 1. Model Performance Comparison

The Random Forest classifier clearly outperformed the Isolation Forest across all metrics, which is expected given its supervised nature and access to labelled data. However, the Isolation Forest's performance is noteworthy, considering it is an unsupervised method operating without prior knowledge of threats.

Figure 2 shows the ROC curves for Isolation Forest and Random Forest models. These curves demonstrate that both models performed significantly better than just random guessing, with the Random Forest showing superior performance in distinguishing between normal and anomalous behaviour across the different classification thresholds.

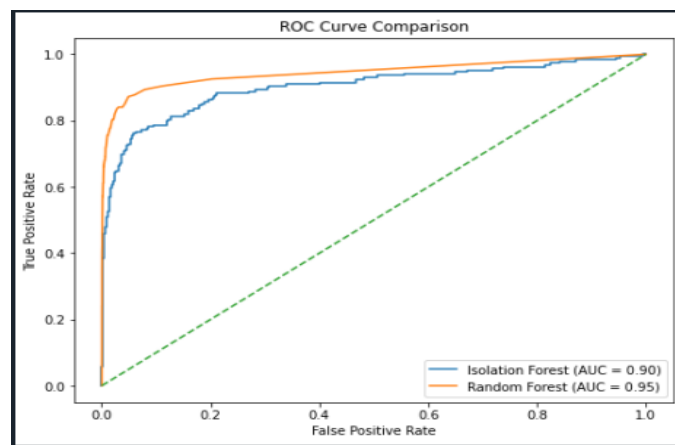


Figure 2. ROC Curve Comparison

### 5.2 Feature Importance

Figure 3 shows the feature importance in the Random Forest model for insider threat detection. It shows the most influential features in detecting insider threats, namely:

- Number of connections (38.66%)
- Bytes transferred (36.83%)
- Duration (4.01%)
- Day of week (3.87%)
- Access level (3.04%)

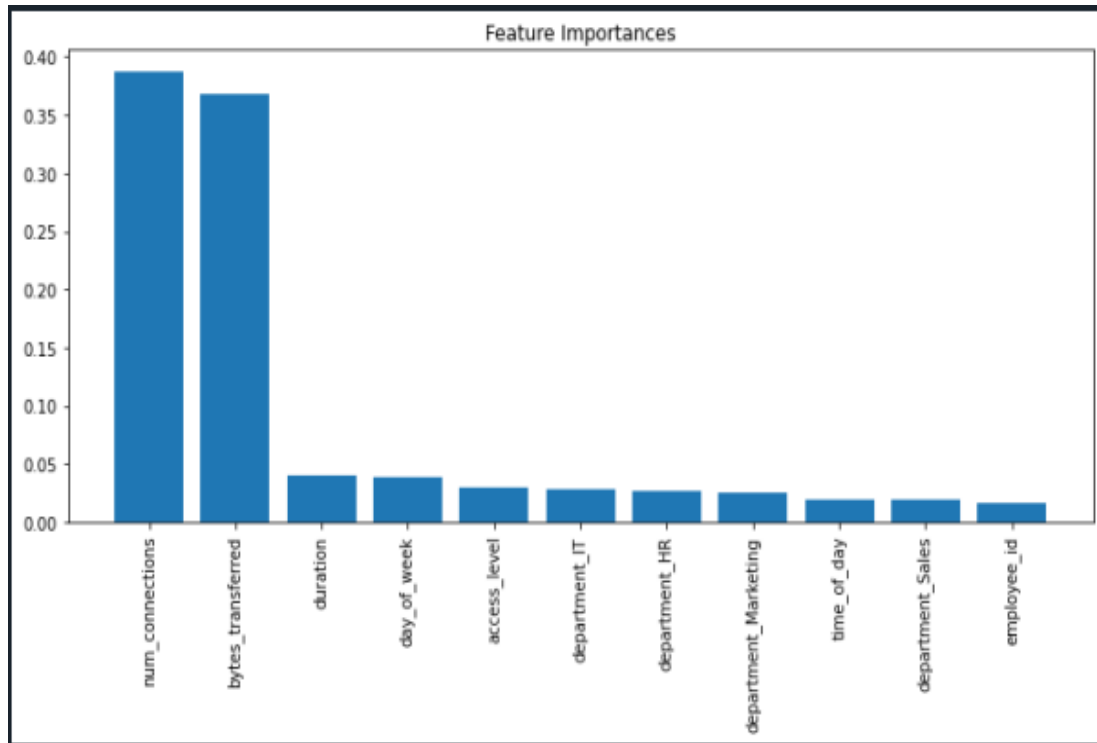


Figure 3. Feature Importance in the Random Forest Model

These findings reveal that frequency and volume of network activity, along with temporal patterns, are crucial indicators of potential insider threats.

### 5.3 Anomaly Detection

Both models identified several instances of anomalous network behaviour. Characteristics of these detected anomalies include:

- High data transfer volumes (ranging from 7,434 to 48,261 bytes)
- Unusual access times (midnight and evening hours)
- Varying connection frequencies (6 to 16 connections)
- Activities spread across different departments

Figure 4 shows the distribution of anomaly scores for the Isolation Forest model. The scores show how the Isolation Forest distinguishes between normal and anomalous behaviour. The long tail on the left represents the data points that the model considers most anomalous. Figures 5 and 6 show the Confusion matrices for the Isolation Forest and Random Forest models, respectively.

The confusion matrices in reveal that while the Random Forest model has a lower false positive rate, the Isolation Forest is more sensitive in detecting potential threats, albeit with more false positives. The models show 73.33% agreement in anomaly classification, highlighting their consistency and potential benefits of combining them.

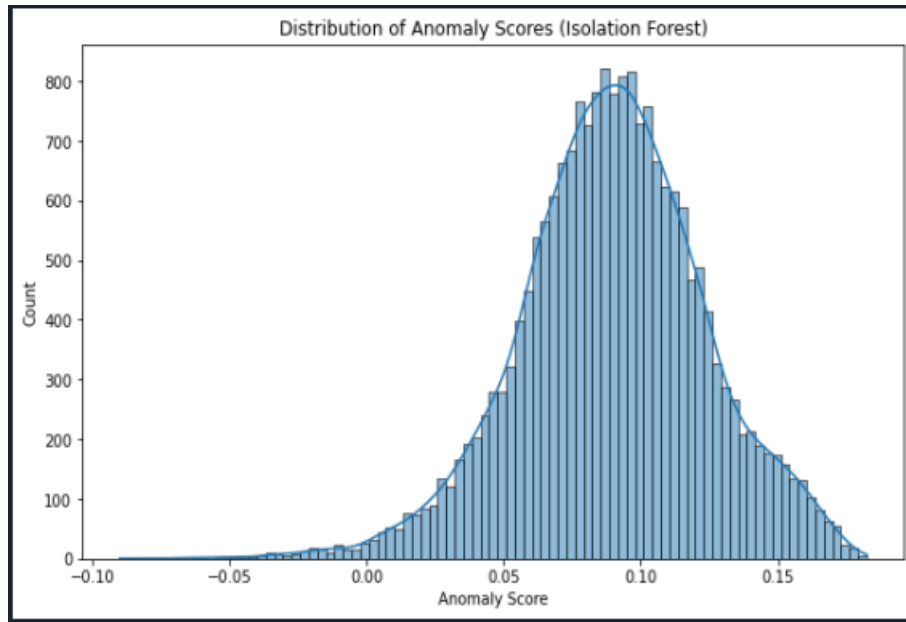


Figure 4. Anomaly Score Distribution

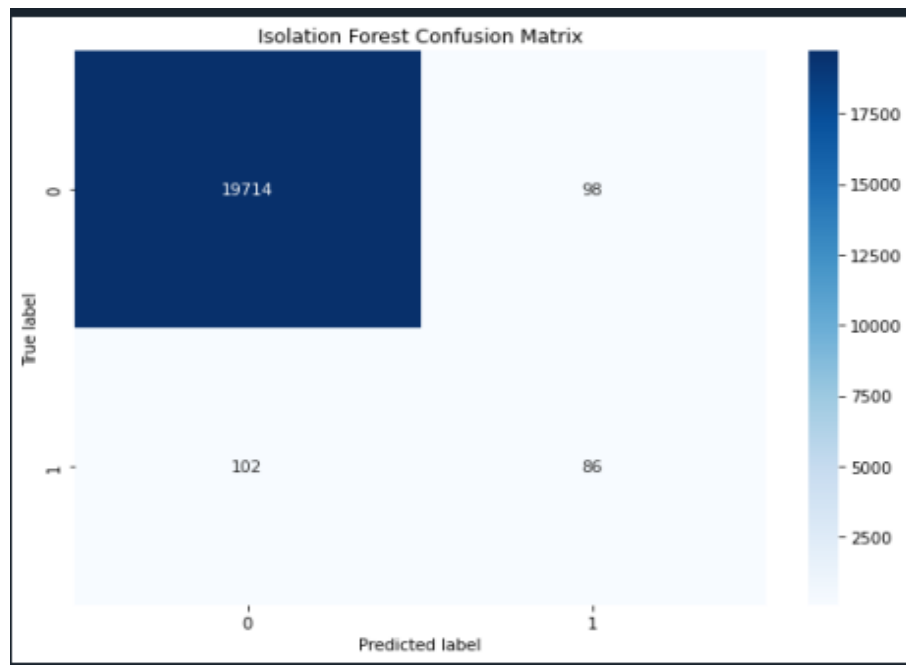


Figure 5. Isolation Forest Confusion Matrix



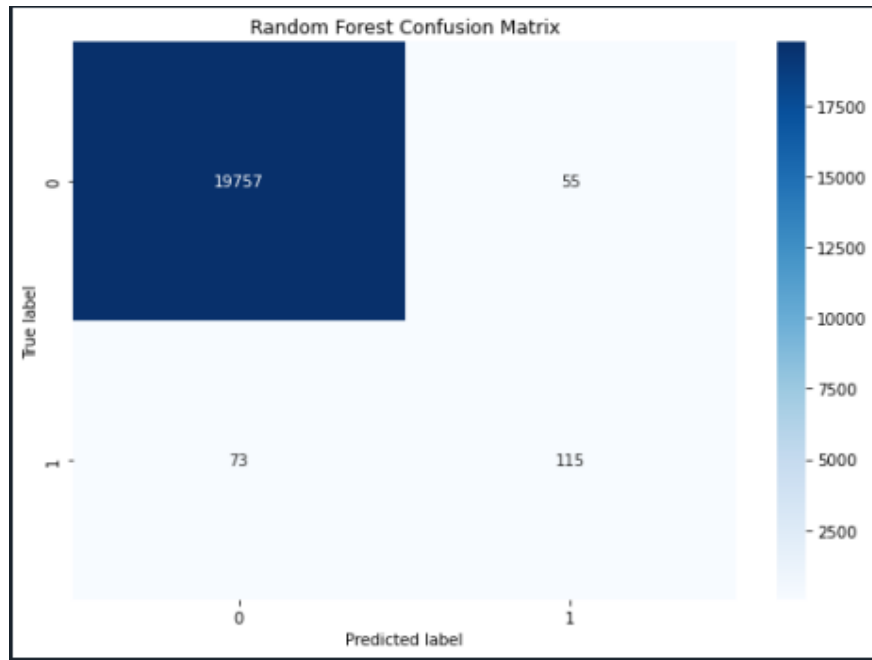


Figure 6. Random Forest Confusion Matrix

#### 5.4 Model Comparison and Practical Implications

The strengths of the unsupervised and the supervised models are shown in Table 2.

Table 2. Strengths of the Models

Isolation Forest	Random Forest
<ul style="list-style-type: none"> <li>• Effective at detecting novel or rare anomalies</li> <li>• Does not require labelled data</li> <li>• Computationally efficient</li> </ul>	<ul style="list-style-type: none"> <li>• Higher overall accuracy</li> <li>• Provides probability scores for more nuanced insights</li> <li>• Allows for feature importance analysis</li> </ul>

The ability of the Isolation Forest model to detect anomalies without prior labelling makes it valuable for identifying new threat patterns. Meanwhile, the higher accuracy of the Random Forest model is beneficial when labelled data is available. These findings align with [3], showing that ensemble methods like Random Forest often outperform single-model approaches. Additionally, the results emphasize the importance of feature engineering, supporting the findings of [7]. These findings align with [3], showing that ensemble methods like Random Forest often outperform single-model approaches. Recent comparative analyses of machine learning models for network anomaly detection further support these conclusions [16], [17].

## 6. IMPLICATIONS FOR DIGITAL FORENSICS

The findings from this work have significant implications for cybersecurity practices, as the combination of unsupervised and supervised methods could provide a more robust defence against insider threats, while prioritizing the monitoring of network connection patterns and data transfer volumes in the detection systems. Contextual factors, such as time of day and department, play a crucial role in identifying suspicious activities. The high agreement rate between models (73.33%) suggests potential for developing hybrid approaches that leverage the strengths of both

supervised and unsupervised techniques, potentially leading to more comprehensive and adaptive insider threat detection systems.

The feature importance analysis provides a data-driven approach, focusing on high-impact features like connection frequency and data transfer volumes. The importance of temporal features highlights the need for tools that can analyse time-based patterns in network traffic. A feedback loop between forensic investigations and the models creates a continuous improvement cycle.

The integration of machine learning in digital forensics aligns with recent trends, as discussed in [12], emphasizing the combination of advanced analytics and human expertise in addressing the complex challenge of insider threats. These machine learning models can be integrated into existing forensic tools for automated triage, improving efficiency and effectiveness.

## 7. LIMITATIONS AND FUTURE WORK

While providing valuable insights into insider threat detection using machine learning, this study does have several limitations. Our use of simulated network traffic data may not fully capture real-world complexity, and the feature set may not encompass all relevant aspects of insider threat behaviour. The focus on Isolation Forest and Random Forest algorithms limits model comparison scope, and the study does not fully address temporal dynamics of evolving threats or consider adversarial scenarios and ethical implications of employee monitoring.

Future work should address these limitations and explore new research avenues. This includes validating models on authentic network logs, expanding the feature set to include user behavioural biometrics and sentiment analysis, and exploring deep learning techniques [18]. This includes validating models on authentic network logs, expanding the feature set to include user behavioural biometrics and sentiment analysis, and exploring deep learning techniques. Developing standardized datasets could facilitate robust comparisons of different methods. Incorporating time series analysis, researching model robustness against adversarial attacks, improving explainability, and addressing privacy considerations through privacy-preserving techniques are crucial. Additionally, investigating real-time implementation challenges, exploring transfer learning applications, and developing hybrid approaches combining supervised and unsupervised methods could lead to more robust and adaptable systems. These research directions aim to contribute to the development of more effective, ethical, and practical insider threat detection systems, enhancing organizational cybersecurity in an increasingly complex digital landscape.

## 8. CONCLUSION

This study compares unsupervised (Isolation Forest) and supervised (Random Forest) machine learning approaches for insider threat detection in network traffic data. The Random Forest model outperformed Isolation Forest, highlighting the value of labelled data, while Isolation Forest showed promise in detecting novel threats. Connection frequency and data transfer volume emerged as critical features for threat detection.

The 73.33% agreement between models suggests potential for hybrid approaches leveraging both techniques. These findings have significant implications for cybersecurity research and practice, demonstrating the potential of data-driven approaches to enhance existing security measures.

While integration of these models into digital forensics could improve insider threat investigations, limitations such as the use of simulated data necessitate further real-world validation. Future research should explore advanced techniques like deep learning and privacy-preserving analytics.

As insider threats evolve, our study contributes to the development of sophisticated, adaptive, and ethical detection systems [6], [19], [20], demonstrating the potential of machine learning in enhancing insider threat detection capabilities, a direction supported by recent reviews in this field [21], [22].

## ACKNOWLEDGEMENT

We thank the anonymous reviewers for the careful review of our manuscript.

## FUNDING STATEMENT

The authors received no funding from any party for the research and publication of this article.

## AUTHOR CONTRIBUTIONS

Sellappan Palaniappan: Conceptualization, Data Curation, Methodology, Validation, Writing – Original Draft Preparation;

Rajasvaran Logeswaran: Project Administration, Writing – Review, Editing & Formatting;

Shapla Khanam: Writing – Review.

## CONFLICT OF INTERESTS

No conflict of interests were disclosed.

## ETHICS STATEMENTS



Our publication ethics follow The Committee of Publication Ethics (COPE) guideline. <https://publicationethics.org/>

## REFERENCES

- [1] Gurukul, "2023 Insider Threat Report," Cybersecurity Insiders, 2023. [\*Online]. Available: [https://library.cyentia.com/report/report\\_014103.html](https://library.cyentia.com/report/report_014103.html)
- [2] P. Chattopadhyay, L. Wang, and Y. P. Tan, "Scenario-based insider threat detection from cyber activities," *IEEE Trans. Comput. Soc. Syst.*, vol. 5, no. 3, pp. 660-675, 2018. doi: 10.1109/TCSS.2018.2857473.
- [3] Z. Azam, M. M. Islam, and M. N. Huda, "Comparative Analysis of Intrusion Detection Systems and Machine Learning-Based Model Analysis Through Decision Tree," *IEEE Access*, vol. 11, pp. 80348-80391, 2023. doi: 10.1109/ACCESS.2023.3296444.
- [4] G. Apruzzese, M. Colajanni, L. Ferretti, A. Guido, and M. Marchetti, "On the effectiveness of machine and deep learning for cyber security," in *Proc. 10th Int. Conf. Cyber Conflict (CyCon)*, 2018, pp. 371-390. doi: 10.23919/CYCON.2018.8405026.
- [5] Y. Xin et al., "Machine learning and deep learning methods for cybersecurity," *IEEE Access*, vol. 6, pp. 35365-35381, 2018. doi: 10.1109/ACCESS.2018.2836950.
- [6] Proofpoint, "Cost of Insider Threats Global Report," Ponemon Institute, 2022. [\*Online]. Available: <https://www.novipros.com/blog/2022-ponemon-cost-of-insider-threats-global-report>
- [7] H. Rai, J. Yoo, and S. Agarwal, "The improved network intrusion detection techniques using the feature engineering approach with Boosting Classifiers," *Mathematics*, vol. 12, no. 24, p. 3909, 2024. doi: 10.3390/math12243909.
- [8] Tuor, S. Kaplan, B. Hutchinson, N. Nichols, and S. Robinson, "Deep learning for unsupervised insider threat detection in structured cybersecurity data streams," *The AAAI Workshop on Artificial Intelligence for Cyber Security*, 2017. doi: 10.48550/arXiv.1710.00811.
- [9] L. Chen et al., "Insider threat detection using Deep Neural Networks with Attention Mechanism," *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 3, pp. 1695-1707, 2022.
- [10] M. Lakshmi et al., "Evaluating the Isolation Forest Method for anomaly detection in software-defined networking security," *J. Electr. Syst.*, vol. 19, no. 4, pp. 279-297, 2023. doi: 10.52783/jes.639.
- [11] S. Ness et al., "Anomaly Detection in Network Traffic Using Advanced Machine Learning Techniques," *IEEE Access*, pp. 1-1, 2025. doi: 10.1109/ACCESS.2025.3526988.

- [12] A. Nisioti, G. Loukas, A. Laszka, and E. Panaousis, "Data-Driven decision support for optimizing cyber forensic investigations," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 2397-2412, 2021. doi: 10.1109/TIFS.2021.3054966.
- [13] A. Shamshari and H. Najaf, "Machine learning approaches for anomaly detection in network security," *Eastern Eur. J. Multidiscip. Res.*, vol. 1, no. 1, pp. 22-29, 2024.
- [14] E. F. Agyemang, "Anomaly detection using unsupervised machine learning algorithms: A simulation study," *Sci. Afr.*, vol. 26, e02386, 2024. doi: 10.1016/j.sciaf.2024.e02386.
- [15] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001. doi: 10.1023/A:1010933404324.
- [16] R. Pandey, M. Pandey, and A. Nazarov, "Advancing Network Anomaly Detection: Comparative Analysis of Machine Learning Models," in *Cryptology and Network Security with Machine Learning*, A. Chaturvedi et al., Eds., Springer, vol. 918, 2024. doi: 10.1007/978-981-97-0641-9\_41.
- [17] M. C. Aswathy, and T. Rajkumar, "Real Time Anomaly Detection in Network Traffic: A Comparative Analysis of Machine Learning Algorithms," *Int. Res. J. Adv. Eng. Hub*, vol. 2, pp. 1968-1977, 2024. doi: 10.47392/IRJAEH.2024.0269.
- [18] L. I. Khalaf et al., "Deep Learning-Based Anomaly Detection in Network Traffic for Cyber Threat Identification," in *Proc. Cognitive Models and Artificial Intelligence Conf. (AICCONF '24)*, pp. 303-309, 2024. doi: 10.1145/3660853.3660932.
- [19] V. P. M. Vishnu Priya and S. Soumya, "Advancements in Anomaly Detection Techniques in Network Traffic: The Role of Artificial Intelligence and Machine Learning," *J. Sci. Res. Technol.*, vol. 2, no. 6, pp. 38-48, 2024. doi: 10.61808/jsrt114.
- [20] P. Gupta and P. Tripathy, "Unsupervised Learning for Real-Time Data Anomaly Detection: A Comprehensive Approach," *Int. J. Comput. Sci. Eng.*, vol. 11, pp. 1-11, 2024. doi: 10.14445/23488387/IJCSE-V11I10P101.
- [21] R. Liu, J. Shi, X. Chen, and C. Lu, "Network anomaly detection and security defense technology based on machine learning: A review," *Comput. Electr. Eng.*, vol. 119, p. 109581, 2024. doi: 10.1016/j.compeleceng.2024.109581.
- [22] J. Ahmad and A. W. Khan, "Empirical investigation of security awareness and training for distributed teams to safeguard from cyber attacks," in *Computing and Data Science*, S.-C. Haw, L. Lee, M. M. Alam, A. Khan, M. Z. Asghar, and F. U. Khan, Eds. MMU Press, 2024, pp. 63-75.

## BIOGRAPHIES OF AUTHORS

	<p><b>Prof. Dr. Sellappan Palaniappan</b> is a Professor of Information Technology at HELP University, Malaysia. With over 30 years of academic experience, his current research interests are in the application of artificial intelligence and machine learning in diverse domains like cybersecurity, data analytics, healthcare, biotechnology, retail, education, agriculture, logistics, and sustainable development initiatives. He is also interested in quantum physics, DNA, neuroscience, and energy, frequency, and vibration for health and wholeness. He has published more than 100 scholarly research papers and authored several widely used IT textbooks for college and university students. He may be contacted at <a href="mailto:sellappan.p@help.edu.my">sellappan.p@help.edu.my</a>.</p>
	<p><b>Prof. Dr. Rajasvaran Logeswaran, SMIEEE</b> is a Professor and Dean of Computing and Digital Technology at HELP University, Malaysia. With over 25 years of academic experience, his research interests are in medical image processing, artificial intelligence, data science and cybersecurity, with over 180 publications in books, peer-reviewed journals and international conference proceedings. He actively serves as a speaker at many international conferences, as well as volunteers as a judge in STEM and innovation competitions for schools at the local and international levels. He may be contacted at <a href="mailto:logeswaran.nr@help.edu.my">logeswaran.nr@help.edu.my</a>.</p>



**Dr Shapla Khanam** is a Senior Lecturer at HELP University, Malaysia. She obtained her PhD from the University of Malaya, where she was also a researcher and helped in the development of Cybersecurity solutions using Deep Learning and Machine Learning for the Internet of Things (IoT). Her research interests include Intrusion Detection, Network Security, Machine Learning, Deep Learning, IoT and WSNs. She has published several articles in journals and presented her findings at international conferences. She may be contacted at [shapla.k@help.edu.my](mailto:shapla.k@help.edu.my).