Behavioural Biometrics and Continuous Authentication for Insider Threat Detection in Enterprise Networks

Nursultan Kuldeyev , Orken Mamyrbayev * , Ainur Akhmediyarova * , Assel Yerzhan

Posted Date: 24 April 2026

doi: 10.20944/preprints202604.1711.v1

Keywords: insider threat detection; behavioural biometrics; continuous authentication; deep learning; LSTM; autoencoder; behavioural risk analysis

 doi:10.20944/preprints202604.1711.v1

Article

Behavioural Biometrics and Continuous Authentication for Insider Threat Detection in Enterprise Networks

Nursultan Kuldeyev [1], Orken Mamyrbayev [2,*], Ainur Akhmediyarova [1,*] and Assel Yerzhan [1]

[1] Satbayev University, Almaty, Kazakhstan

[2] Institute of Information and Computational Technologies, Almaty, Kazakhstan

* Correspondence: morkenj@mail.ru (O.M.); a.akhmediyarova@satbayev.university (A.A.)

Abstract

Identifying insider threats in modern enterprise environments presents a unique cybersecurity challenge. Although malicious activity may often appear to be legitimate user activity, it is difficult to recognize the distinction. This study presents an innovative approach to insider threat detection by analyzing enterprise activity logs for continuous authentication along with behavioural biometrics. Behavioural patterns, such as logins, file accesses, network interactions and emails, are analyzed to determine abnormal behaviours of users. The proposed system utilizes a hybrid deep learning architecture that includes a Long Short-Term Memory (LSTM) network and an autoencoder model to model temporal dependence of a user's behaviour and to identify anomalies through reconstruction error analysis. The LSTM network captures user's sequential activity and autoencoder determines variance from the user's typical behavioural profile. The outputs of both models are aggregated using a unified behavioural risk scoring mechanism for continuous authentication and an ongoing assessment of insider threats. The experimental results from Insider Threat Dataset for Corporate Environments demonstrate that proposed approach is effective in classifying normal versus malicious behaviours of users. The model achieves of 97.65% an accuracy, of 96.35% a precision, of 99.05% a recall rate, of 97.68% an F1-score and a Receiver Operating Characteristic - Area Under Curve (ROC-AUC) score of 99.20%, which indicates a high level of detection capability and very low false positives. The findings support that a developed model is a viable solution for integrating behavioural modelling, detection of anomalies.

Keywords: insider threat detection; behavioural biometrics; continuous authentication; deep learning; LSTM; autoencoder; behavioural risk analysis

1. Introduction

Digital transformation has altered how businesses function at a rapid pace; as a result, many businesses now depend on multiple interrelated systems, cloud infrastructures, and vast amounts of data [1]. Technology improves business performance (efficiency and productivity), but the technology also creates many complex challenges related to cybersecurity [2]. Insider threat can be recognized as one of the most significant types of threat to organisations because insider threats can be generated by someone with legitimate access to organisational systems [3]. In contrast to external attacks, insider threats tend to be difficult to detect because the actions of authorised users are often indistinguishable from normal activities [4]. Enterprises produce large volumes of logs that contain records of logins, file access, email records, and network browsing, which offer valuable insight into how users behave in the enterprise; therefore, advanced analytic solutions need to be used to monitor user behaviours in order to provide timely identification of potential threats to a business's security [5].

 doi:10.20944/preprints202604.1711.v1

There can be many reasons for insider threats which come from a variety of organizational, behavioural and technical influences affecting users' actions on an enterprise system [6]. The most common reason why malicious insider threat occurs is because insiders exploit their privileged access privileges to steal sensitive information, sabotage organizational resources, or gain financial gain [7]. There are also other times when insider threats occur because of human errors or lack of cybersecurity awareness [8] (unintentional or accidental misuse of an organization's resources). Some of these contributing factors are weak access control mechanisms, excessive privilege allocations, limited monitoring of user activity and credential compromised by phishing and social engineering type attacks [9]. Because of the increased prevalence of remote work environments, along with distributed enterprise networks, the difficulty in monitoring employee behavior has increased for organizations to be able to detect abnormal or unauthorized access patterns in real-time.

Existing studies have proposed various technical solutions for identifying insider threats and enhancing enterprise security [10]. Traditional security solutions are based on rules, known as rules-based monitoring solutions; intrusion detection systems are based on signatures, also known as signature-based intrusion detection systems [11]; and the other type of traditional security solution is based on static user identification, also known as static authentication methods. These traditional security solutions do not typically offer the capabilities needed to identify insider threats because it can be difficult to identify; as a result, these solutions do not work for sophisticated insider threats because malicious activity may occur outside of known malicious patterns [12]. In order to address this issue, a number of studies have begun to look at the application of machine (ML) and deep (DL) learning algorithms to analyze large volumes of behaviour data collected over time within enterprise environments [13]. Sequential learning algorithms, such as (LSTM) models, have been shown to effectively learn from the temporal dependencies of user activity, which allows them to identify abnormal behaviour over time. Likewise, reconstruction-based neural networks (e.g., Autoencoders) can be used to model normal behaviours and identify behaviours that deviate from this norm and indicate possible insider threats.

This paper suggests a mechanism of continuous authentication for detection of insider threats in a business to improve upon traditional detection methods. The proposed method will look at the user's behaviour patterns based off the enterprise activity logs used when looking at the CERT Insider Threat Dataset. Deep Learning Techniques will be used in the Framework for the purposes of sequential behaviour modelling and Group Anomaly Detection. The proposed framework will continuously examine the individual behaviour patterns of users and compare them against the known behaviour profiles of those users. By using Temporal Sequence Learning, the proposed framework will be able to assess and collect the changing behaviour over time, and by employing Reconstruction Based Anomaly Detection the Framework can find a small but significant change in behaviour from what is determined to be "normal." Through continuous monitoring of behaviours and risk scoring methodologies, the Framework will assist in providing more accurate detection of insider threats while offering a more adaptive and proactive security solution in today's modern enterprise environments. Contributions of the Study,

- ➔ Proposed behaviour-based insider threat detection model is an approach to rate a pattern of the user activity and detect any abnormal behaviour within enterprise environment.
- ➔ Hybrid deep learning architecture Hybrid deep learning architecture combining the Autoencoder with an LSTM network to capture both temporal dependencies and anomalies in normal user behaviour.
- ➔ The dynamic behavioural risk scoring mechanism is presented to assess the probability of insider threats throughout user sessions.
- ➔ Behavioural analysis is done to study the risk score distribution, time series development of insider threat risk, and user authentication stability.
- ➔ Experimental test on CERT Insider Threat Data set on Corporate Environments show that insider threat detection can be done successfully with high accuracy and minimized false positives.

 doi:10.20944/preprints202604.1711.v1

1.1. Research Organization

The remainder of this paper is organized as follows. Section 2 presents a review of related studies on insider threat detection and behavioural analysis techniques used in enterprise security environments. The proposed methodology in Section 3 contains system architecture and data preprocessing procedures and feature extraction process and LSTM–Autoencoder model development for behavioural anomaly detection. The experimental setup and dataset description, which includes preprocessing steps for Insider Threat Dataset for Corporate Environments, are detailed in Section 4. The results and performance evaluation, which includes behavioural risk score analysis and temporal risk evolution and session authentication stability and confusion matrix evaluation and overall performance metrics of the proposed LSTM–Autoencoder Framework of Section 5 presents. The study ends with Section 6, which presents research directions for future work to enhance insider threat detection systems.

2. Literature Review

M. S. Mohamed and A. Arabo [14] proposed a Security Information and Event Management (SIEM) integrated framework for insider threat anomaly detection using enterprise logs and behavioural biometrics. The system analyses user behavioural patterns extracted from organizational log data to identify suspicious activities. Their prototype combines behavioural analytics with SIEM monitoring to improve real-time detection capabilities. Experimental results demonstrate improved anomaly detection performance in enterprise environments. The study highlights the importance of integrating behavioural biometrics with centralized security monitoring systems. S. S. P. Pennada et al. [15] investigated insider threat detection using behavioural analysis supported by machine learning and deep learning models. Their approach analyses user activity patterns such as login behavior, file access, and system interactions to detect anomalous activities. The study evaluates multiple algorithms to determine the most effective models for insider threat identification. Results show that deep learning techniques outperform traditional machine learning methods in identifying complex behavioural anomalies. The research emphasizes the effectiveness of behavior-based monitoring for enterprise security.

S. S. Abba et al. [16] introduced a behavioural biometrics-powered continuous authentication framework designed for zero-trust remote work environments. The proposed system combines multiple identity verification mechanisms including behavioural patterns and biometric indicators to strengthen authentication processes. The framework supports continuous monitoring of user behavior to detect unauthorized access attempts. Their results demonstrate improved identity verification accuracy in distributed working environments. The study highlights the importance of multi-factor behavioural authentication for secure remote systems. O. O. Aramide [17] examined the application of artificial intelligence techniques for identity verification and authentication in network security systems. The study integrates biometric identification and behavioural analytics to enhance authentication accuracy and security efficiency. AI-driven models are used to analyse user behavior patterns and detect anomalies in network interactions. The results indicate that AI-based authentication systems improve both speed and accuracy compared with traditional verification mechanisms. The research supports the use of intelligent authentication systems in modern cybersecurity infrastructures.

I. Ibraheem et al. [18] developed an anomaly-based user behavior analytics framework which detects insider threats in commercial networks. The system uses machine learning algorithms to analyze activity logs and detect insider attack patterns which show unusual behavior. The system detects suspicious user activities through monitoring users who show unusual access patterns and unexpected network usage. The results of the experiment show that behavioral anomaly detection brings major improvements to threat detection success rates. The study confirmed that user behavior analytics plays a vital role in protecting enterprise systems from security threats. J. Hu et al. [19] developed an enterprise internal threat authentication traceability technology based on a key

 doi:10.20944/preprints202604.1711.v1

authentication system. The proposed approach focuses on tracking user authentication processes to ensure traceability and accountability within enterprise systems. The framework provides secure identity verification mechanisms to prevent unauthorized access and internal misuse. Experimental evaluations demonstrate improved reliability and security in authentication processes. The research highlights the significance of traceable authentication mechanisms for mitigating internal threats.

D. He et al. [20] presented a two-level deep learning model to identify internal threats in enterprise systems. The model considers two detection layers to process user behavior pattern and system activity logs. Deep learning algorithms are used to learn complicated associations between the user behavior and possible signs of threats. According to the experimental results, there are better detection accuracies than conventional anomaly detection methods. The paper proves that deep learning architectures can be used to detect advanced forms of insider threats. K. Saminathan et al. [21] introduced an artificial neural network autoencoder model with insider cybersecurity threats detection. The method proposed works by obtaining the normal patterns of user behavior through system logs and detecting the abnormal ones by analysing reconstruction errors. Anomaly detection using autoencoders facilitates the detection of suspicious user behavior, and does not need a large amount of labelled data. According to experimental assessment, high presentation in the identification of insider threat patterns is demonstrated. The study shows that deep learning-based anomaly detection methods are useful in cybersecurity applications.

U. Uslu et al. [22] performed a comparative assessment of deep learning structures to continuous authentication by behavioural biometrics. The interaction patterns among the users including typing behavior and characteristics of system use are analyzed in the study to authenticate the users at all times. Different deep learning structures were tried to understand the ability of the architectures to capture behavioural features. It has been found out that deep learning models are characterized by a high level of authentication accuracy and reliability. According to the research, behavioural biometrics has the capacity to enhance authentication systems. A. Orun et al. [23] researched cognitive behavioural features in remote user authentication in cybersecurity systems. The offered solution examines behavioural characteristics connected to the way the interaction occurs between the users to enhance identity verification procedures. Their model is based on the determination of distinctive behavioural features characterizing legitimate users and attackers. The experimental findings indicate that there is enhanced reliability in authentication in remote access set ups and that security risks are limited. The study identifies the relevance of cognitive behavioural analysis in the contemporary authentication systems. X. Tao et al. [24] presented a system of insider user authentication, which is built upon a better temporal convolutional network. The model identifies abnormal activity of authentications by capturing temporal relationships among series of user activities. The enhanced structure boosts feature dispensing of successive user communication information. Detection accuracy is found to be high through experimental assessments as opposed to traditional authentication processes. This research validates that this could prove beneficial for detecting insider threats and in systems of continuous authentication.

2.1. Problem Statement

The detection of insider threats within the enterprise workplace remains challenging due to advancements in cybersecurity technologies, the complexity of user behaviour, and the instability of malicious behaviour [25]. The existing security solutions rely primarily on traditional rule-based methods (e.g., rule-based monitoring systems, signature-based intrusion detection systems, static authentication) that rely heavily on pre-defined rules or just validating a user's credentials [26,29]. These traditional security solutions are ineffective for detecting sophisticated insider attacks because they allow an authorized user to conduct malicious actions, which will appear similar to normal operations and thus will be undetectable by these traditional methods [27]. Additionally, the growing amount of data being processed in the enterprise and increasing amounts of data being used in remote working environments combined with the growing complexity of enterprise networks have made it more difficult for organizations to conduct continuous monitoring of user activity and detect

 doi:10.20944/preprints202604.1711.v1

anomalies in real-time. Therefore, a proposed approach to overcome the challenges associated with detecting and mitigating insider threats is to develop a continuous authentication framework for insider threats that uses a behavioural biometrics approach to identify anomalies utilising user activity logs from the data through the application of deep learning models (e.g., LSTM, Autoencoder) that will capture temporal behaviours of users to detect anomalies and thus increase the accuracy and reliability of insider threat detection in enterprise networks.

3. Methodology

The proposed framework presents the **Figure 1** behavioural biometrics–driven approach for continuous authentication and insider threat detection in enterprise environments. The process begins with collection of enterprise activity logs which include insider threat dataset login events and file access and network interactions. The system undergoes a data preprocessing stage which performs cleaning and normalization to ensure data quality and consistency. User activity pattern indicators are developed through the process of behavioural feature extraction. The LSTM model analyzes behavioural sequences because it learns how users behave over time through its ability to capture temporal dependencies. The autoencoder model identifies abnormal user behaviour through its behaviour reconstruction process which uses reconstruction error as an anomaly indicator. The system integrates LSTM temporal modelling outputs with autoencoder reconstruction results to generate an insider threat score which measures both temporal behavioural changes and reconstruction anomalies. The risk score enables a continuous authentication mechanism which monitors user behaviour through multiple windows to identify normal activities and potential insider threats. The proposed framework achieves its effectiveness evaluation through multiple performance metrics which include accuracy and precision and recall and F1-score and confusion matrix analysis and behavioural risk analysis through risk score distribution and aggregated insider risk trends and temporal risk evolution and session authentication stability and behavioural stability comparison. The proposed system demonstrates reliability for proactive insider threat detection through its multiple evaluation methods which include accuracy and precision and recall and F1-score and confusion matrix analysis and behavioural risk assessment through risk score distribution and aggregated insider risk trends and temporal risk evolution and session authentication stability and behavioural stability comparison.

3.1. Behavioural Feature Extraction

The process of feature extraction identifies essential user behavior patterns which security experts use to detect unusual user activities. The CERT Insider Threat Dataset provides pre-processed logs that researchers use to study various user interaction patterns which occur when users access enterprise systems through their login activities and file access and network usage and email communication.

The behavioural signs which users present to organization enable security professionals to analyse user behavior patterns which help them identify normal activity and potential security threats. The system stores user behavior data in structured feature representations which researchers use to track user behavior over different time periods. The system uses these representations as inputs for sequential learning models which include LSTM and Autoencoder and reconstruction-based learning models to identify patterns of normal behavior which help detect insider threats.

 doi:10.20944/preprints202604.1711.v1

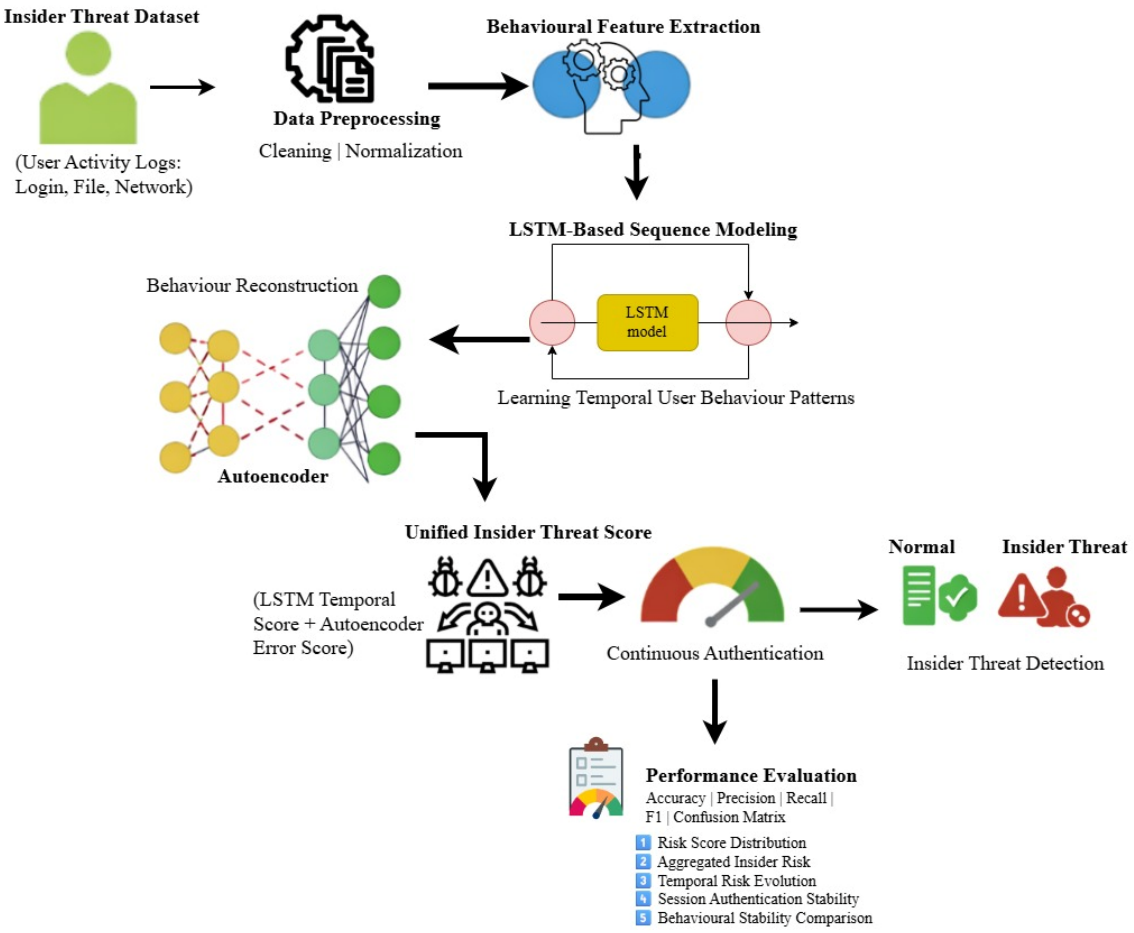


Figure 1. Proposed Behavioural Biometrics-Based Insider Threat Detection Framework.

3.1.1. User Login Pattern Features

The system tracks user access patterns to determine when users access enterprise system and how often they access it. The system tracks three user authentication patterns, which include login frequency and login time distribution and session duration. Login frequency represents how often a user logs into the system within a specific time period. It can be expressed as shown in Equation (1).

$$LF_u = \sum_{i=1}^{n} L_i \quad (1)$$

where LF_u denotes the login frequency of user u, L_i represents an individual login event, and n indicates total number of login events observed during the monitoring period. Higher or unusual login frequencies may indicate suspicious behavior.

3.1.2. File Access Behavior

The system tracks all user activities that show how they interact with organization's documents and files. Abnormal file access activities, such as accessing a large number of files or unusual file types, may indicate insider threats. The file access rate can be computed as shown in Equation (2).

$$FA_u = \frac{N_f}{T} \quad (2)$$

where FA_u denotes the file access rate for user u, N_f represents number of files accessed, and T denotes observation time interval. Significant deviations from normal file access patterns may indicate potential data exfiltration activities.

 doi:10.20944/preprints202604.1711.v1

3.1.3. Network Activity Features

These features represent user interactions resources and network services. The features incorporate data about visited URLs and request frequency and browsing behavior. The network request frequency can be expressed as shown in Equation (3).

$$NR_u = \sum_{j=1}^{m} R_j \quad (3)$$

where NR_u denotes the total number of network requests generated by user $u_r R_j$ represents an individual network request, and m is total number of requests within observation period. Unusual browsing patterns or excessive network requests may indicate suspicious activities.

3.1.4. Email Communication Behavior

It can be tracks the patterns that occur when employees send and receive emails through the enterprise network. Employees who attempt to commit insider threats will display unusual communication patterns which include sending excessive emails and sending large file attachments. Email activity can be quantified as shown in Equation (4).

$$EA_u = E_s + E_r \quad (4)$$

where EA_u denotes total email activity of user u, E_s represents number of sent emails, and E_r represents the number of received emails. Abnormally high email activity or unusual attachment sharing may indicate potential insider threats.

The extracted behavioural features created a complete view of how users interacted with enterprise system. The resulting feature vectors are subsequently used as input to models such as LSTM and Autoencoder to detect anomalous behaviours associated with insider threats.

3.2. Deep Learning-Based Behavioural Sequence Modeling

The analysis of user behavioural patterns needs to include their time-based security track record which enables organizations to identify insider threats that exist within their corporate networks. User activities such as login attempts, file access operations, network browsing, and email communications occur sequentially over time and form behavioural patterns that can reveal abnormal activities. The detection of sophisticated behavioural changes needs more than static analysis because attackers create their actions to look like normal user behavior. The system uses sequential learning techniques to track user activity patterns through time and to detect hidden changes in behavior. The study uses deep learning models to analyze time-series sequences of behavioural feature vectors which enables the system to learn typical user behavior and identify suspicious activities that could indicate insider threats.

3.2.1. LSTM-Based Behavioural Sequence Modeling

The LSTM network serves as tool to model user activity patterns which exhibit changes over time. The LSTM network functions as a specialized type of recurrent neural network which enables sequential data analysis through its ability to maintain extended connections while solving vanishing gradient issue found in standard recurrent networks. The LSTM system uses its internal memory cell function to store vital behavioural data throughout several time periods. This function enables model to improve its ability to identify abnormal user activities by mastering complex user activity patterns.

Figure 2 shows the LSTM architecture which proposed framework uses. The LSTM network functions to analyze user behavior by processing their sequential activity data through its ability to detect time-based relationships. The system uses three main gates which include memory cell uses three gates to govern data movement through its system. The forget gate determines which information from previous cell state should be retained or discarded., while input gate integrates new behavioural information obtained from current user activity features. The output gate generates hidden state that represents the learned behavioural pattern at each time step. The LSTM model in

 doi:10.20944/preprints202604.1711.v1

proposed insider threat detection framework analyzes sequential behavioural features which include login activities and file access patterns and session interactions from the Insider Threat Dataset for Corporate Environments. The model learns time-based patterns of normal user behavior which allows it to detect suspicious behavior that matches insider threats thus enabling continuous authentication and dynamic behavioural risk assessment for enterprise environments.

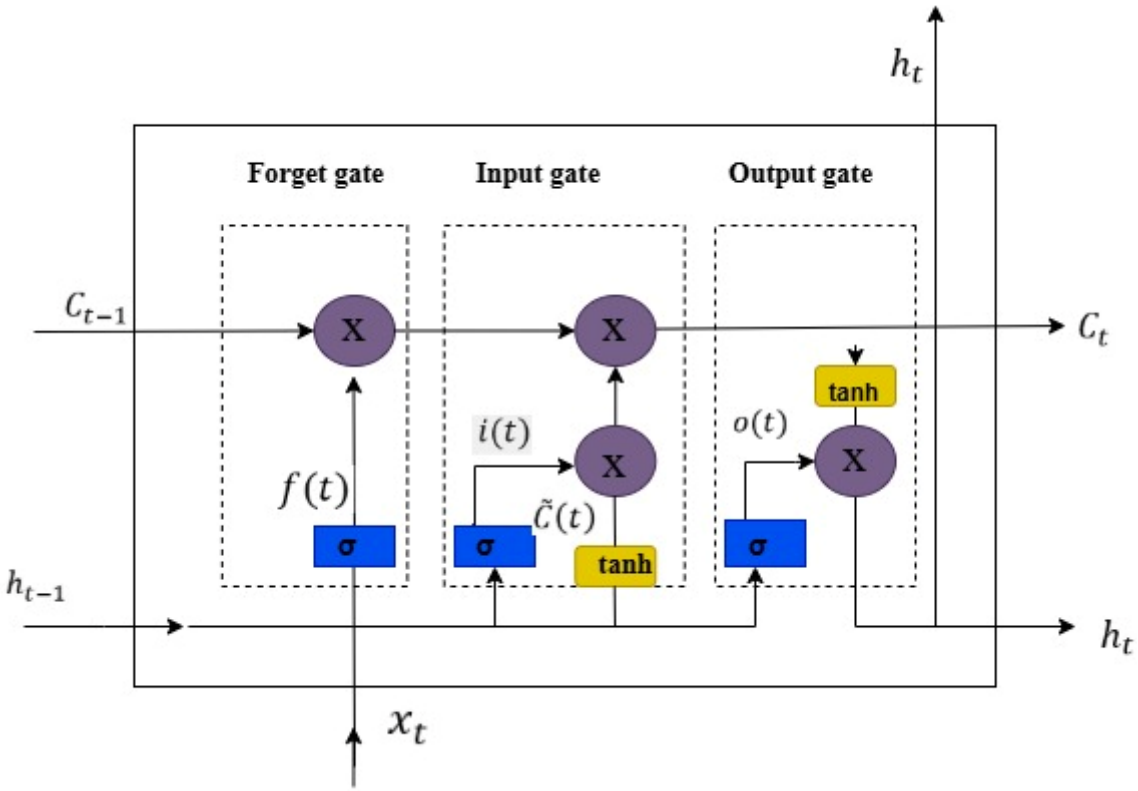


Figure 2. LSTM-Based Behavioral Pattern Modeling for Insider Threat Detection.

The hidden state of LSTM network at time step t is computed as shown in Equation (5).

$$h_t = f(W_h x_t + U_h h_{t-1} + b_h) \tag{5}$$

where h_t represents the hidden state at time step t, x_t denotes input behavioural feature vector at time t, h_{t-1} represents the hidden state from the previous time step, W_h and U_h are weight matrices associated with the input and recurrent connections, b_h denotes the vector of bias, and $f(\cdot)$ represents the activate function.

To regulate the flow of information through the network, LSTM employs gating mechanisms. The input gate controls how much new information should be stored in the memory cell and is defined as shown in Equation (6).

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \tag{6}$$

where i_t represents the input gate activation, W_i and U_i denote the weight matrices, b_i is the bias term, and σ represents the sigmoid activation function.

The 'forget gate' is responsible for the extent of memory information from the previous memory state that needs to be retained and not quashing or forgetting discarded, as shown in Equation (7).

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \tag{7}$$

where f_t denotes the forget gate activation, and W_f, U_f, and b_f represent the corresponding parameters.

 doi:10.20944/preprints202604.1711.v1

The memory cell state is updated by merging the retained and newly learned information., as defined in Equation (8).

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{8}$$

where C_t represents the updated memory cell state, C_{t-1} denotes the previous memory state, and $\tilde{C}_t$ represents the candidate memory content generated by the network.

3.2.2. Autoencoder-Based Anomaly Detection

The system uses an anomaly detection system that operates through a Autoencoder model to detect unusual user behavior in enterprise environments. Autoencoders function as deep learning models that operate without supervision to develop compact data representations which they use to reconstruct entire input systems through their compression and reconstruction process. The researchers trained an autoencoder model with normal user behavior data taken from enterprise activity logs which they collected from the CERT Insider Threat Dataset. The model develops its understanding of user behavior through training which enables it to identify authentic user actions. The system uses reconstruction error measurement to detect potential insider threats when trained model receives new behavioural patterns which differ from its established normal operating patterns.

The insider threat detection system uses an autoencoder architecture which is shown in **Figure 3**. The autoencoder consists of two primary components, namely encoder and decoder, connected through a latent representation layer. The encoder transforms high-dimensional behavioural input data into a compact latent feature space that captures the essential characteristics of user activities. Decode the initial content into the recovered representation which it does by minimizing reconstruction loss during training. The autoencoder uses normal user behavior data which is collected from the Insider Threat Dataset for Corporate Environments to train its system. The model produces higher reconstruction errors when users display abnormal behavior or engage in malicious activities because these actions differ from established behavior patterns. The system uses reconstruction deviations as indicators for detecting unusual behavior which helps to calculate a behavioural risk score within complete detection system. The autoencoder system enhances insider threat detection and continuous authentication systems for enterprise environments by creating compact behavioural models and detecting reconstruction anomalies.

1. Encoder Layer

The encoder network converts high-dimensional behavioural feature vector data into a compact latent representation which maintains essential user behavior attributes. The model achieves its main objective by converting data into fewer dimensions which allow it to concentrate on important behavioural elements while eliminating unnecessary data. The encoding process is mathematically represented as shown in Equation (9).

$$z = f(W_e x + b_e) \tag{9}$$

where z denotes the latent representation, x represents the input behavioral feature vector, W_e denotes the encoder weight matrix, b_e represents the bias vector, and $f(\cdot)$ is the nonlinear activation function applied in the encoder layer.

2. Decoder Layer

The decoder network reconstructs original behavioural feature vector from a latent representation generated by encoder. The decoder needs to produce a complete and exact match of an input data. The system produces inaccurate reconstruction results when the input behavioural pattern differs from normal patterns which the system learned during training. The decoding process is defined as shown in Equation (10).

$$\hat{x} = g(W_d z + b_d) \tag{10}$$

where $\hat{x}$ denotes reconstructed feature vector, z represents latent representation obtained from the encoder, W_d denotes the decoder matrix of weight, b_d represents the bias vector, and $g(\cdot)$ denotes the activation function used in decoder layer.

 doi:10.20944/preprints202604.1711.v1

3. Reconstruction Error

The reconstruction error between original input and reconstructed output is used to check whether or not a behavioural sequence is normal or anomalous. This is the error used as anomaly score in order to signal dangerous user behaviours. The error of reconstruction is calculated as shown in Equation (11).

$$E = \|x - \hat{x}\|^2 \tag{11}$$

where E represents the reconstruction error, x denotes the original behavioral feature vector, and $\hat{x}$ represents reconstructed feature vector produced by autoencoder model. Increase in reconstruction error implies an even greater deviation of normal behavioural patterns and could be an indicator of possible insider threat activities in the enterprise system.

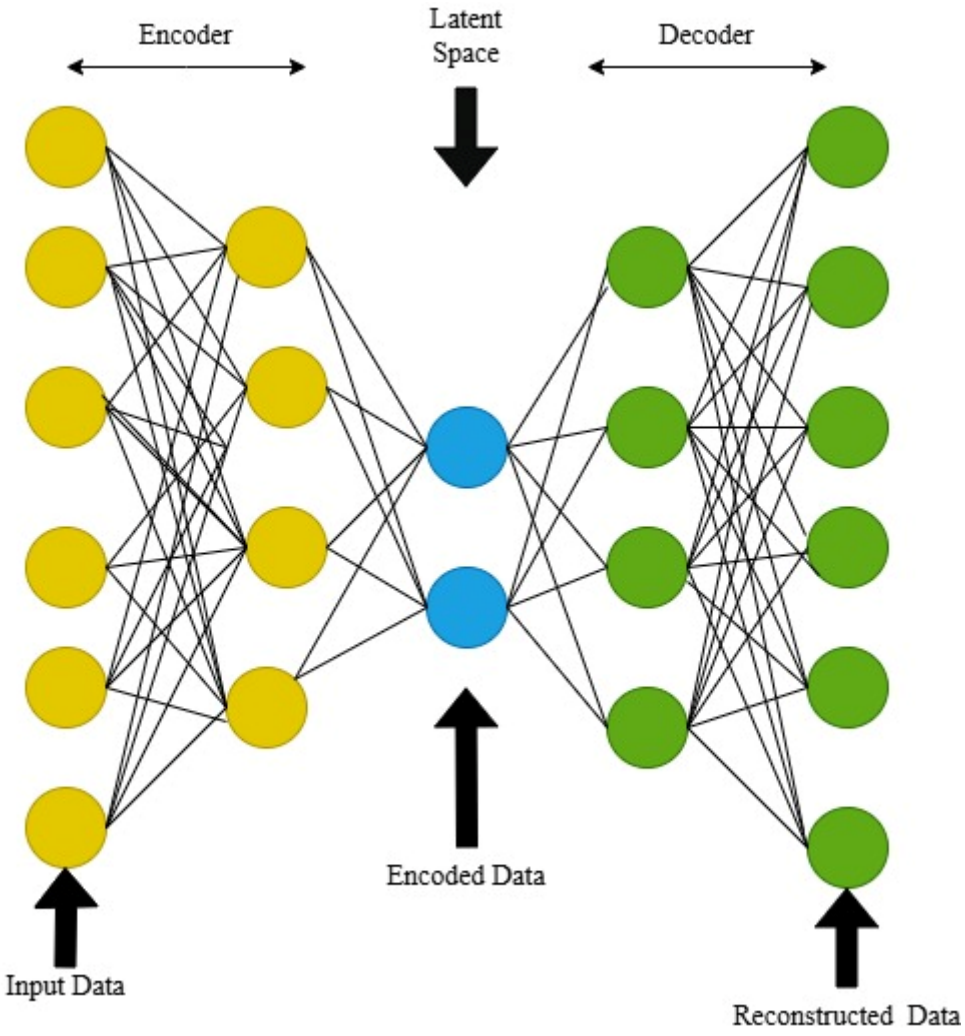


Figure 3. Autoencoder-Based Behavioral Anomaly Detection for Insider Threat Identification.

3.2.3. Unified Insider Threat Score

The combined insider threat score is a combination of both the temporal behavioural deviation and reconstruction-based anomaly scores that produces a powerful and dynamic risk indicator. Because insider threats can frequently be said to be subtle changes in behaviour over time as opposed to a sudden abnormal behaviour, an outlier measure can result in an unstable detection. Hence, suggested architecture is a combination of successive dependency learning as implemented by the LSTM with reconstruction error as implemented an autoencoder.

 doi:10.20944/preprints202604.1711.v1

The temporal anomaly component is derived from LSTM hidden representation represented by Equation (12):

$$S_{LSTM}(t) = g(h_t) \tag{12}$$

where h_t represents the temporal behavioral embedding at time window t, and $g(\cdot)$ is a transformation function (e.g., fully connected layer with sigmoid activation) that converts hidden state into a probability-based anomaly score. This score captures deviations in evolving user behavior patterns across time windows.

The reconstruction-based anomaly component is defined as Equation (13):

$$S_{AE}(t) = \mathcal{L}_{rec}(t) \tag{13}$$

where $\mathcal{L}_{\text{rec}}(t)$ denotes the mean squared reconstruction error between actual behavioral features and reconstructed features. High reconstruction error values demonstrate that data is acting unusually as compared to normal behavioural observations.

The final insider threat score is computed as a weighted combination denoted by Equation (14):

$$S_{\text{final}}(t) = \alpha S_{LSTM}(t) + (1-\alpha) S_{AE}(t) \tag{14}$$

where:

- $\alpha \in [0,1]$ controls the relative importance of temporal learning versus reconstruction deviation.
- When α is higher, the system emphasizes sequential behavioral evolution.
- When α is lower, reconstruction deviation plays a dominant role.

The combination method strengthens detector performance through its balanced approach between predictive modeling and anomaly reconstruction. The system detects false positives which occur from temporary behavioural changes while it maintains detection sensitivity for gradual insider threat development.

To enable continuous authentication, the risk score is dynamically evaluated across consecutive time windows. The classification decision is determined by comparing unified score with a predefined threshold τ shown by Equation (15):

$$\text{Threat}(t) = \begin{cases} 1, & \text{if } S_{\text{final}}(t) > \tau \\ 0, & \text{otherwise} \end{cases} \tag{15}$$

The system identifies a user as high-risk when their score exceeds established threshold during multiple assessment periods. The system allows for early user intervention based on this identification.

Algorithm 1: Behavioural Risk Score Generation Using LSTM–Autoencoder

Input: User activity dataset D

Output: Behavioral risk score R for each user session

1: Load dataset D

2: Preprocess dataset (remove missing values, normalize features)

3: Divide dataset into user sessions S

4: Initialize LSTM–Autoencoder model M

5: Train model M using normal behavioral data

6: For each session s in S do

7: Extract behavioral feature vector F from session s

8: Reconstruct F using trained model M

9: Compute reconstruction deviation between original and reconstructed features

10: Calculate risk score R(s) based on deviation level

 doi:10.20944/preprints202604.1711.v1

11: Store R(s)
12: End For
13: Return behavioral risk scores for all sessions

3.3. Continuous Authentication and Dynamic Risk Assessment Mechanism

The new framework extends existing systems which require one-time authentication to create a continuous system that verifies user identity through their ongoing behavioural patterns. The system continuously checks user behavior through time-based analysis which uses behavioural biometric data from enterprise logs in CERT Insider Threat Dataset to evaluate user activity. This enables both monitoring activities and adjusting risk assessments.

3.3.1. Continuous Authentication Model

Let $S_{\text{final}}(t)$ denote the unified insider threat score at time window t. Instead of making a single decision per session, system computes a dynamic confidence score over multiple consecutive windows.

The cumulative behavioural risk over T windows is defined as Equation (16):

$$R_T = \frac{1}{T}\sum_{t=1}^{T} S_{\text{final}}(t) \tag{16}$$

where:

- $R_T =$ aggregated user risk score
- $T =$ number of monitored time windows
- $S_{\text{final}}(t) =$ insider threat score at window t

Corrections typically smoothen temporary behaviour anomalies, resulting in fewer false alarms coming from single-incident anomalies.

3.3.2. Dynamic Risk Escalation

The implementation of a persistence-based escalation mechanism strengthens security systems by improving their detection accuracy. When anomaly scores surpass the threshold τ during three consecutive time periods, the system detects elevated risk levels represented by Equation (17):

$$E_t = \begin{cases} E_{t-1} + 1, & \text{if } S_{\text{final}}(t) > \tau \\ 0, & \text{otherwise} \end{cases} \tag{17}$$

where:

- $E_t =$ escalation counter
- $\tau =$ anomaly threshold

If $E_t \geq k$ (predefined persistence threshold), the system triggers a security alert. This prevents singlewindow spikes from generating unnecessary alarms.

3.3.3. Continuous Authentication Decision

The final authentication decision is based on the cumulative risk shown by Equation (18):

$$\text{Access Status} = \begin{cases} \text{Trusted}, & \text{if } R_T \leq \gamma \\ \text{High Risk}, & \text{if } R_T > \gamma \end{cases} \tag{18}$$

where γ is the system risk tolerance threshold.

Algorithm 2: Insider Threat Detection and Continuous Authentication
Input: Behavioral risk scores R
Output: Threat classification result C
1: Define risk threshold T

 doi:10.20944/preprints202604.1711.v1


```
2: Initialize classification result set C
3: For each user session i do
4:      Retrieve risk score R(i)
5:      If R(i) > T then
6:          Label session as malicious
7:          Update C(i) = Insider Threat
8:      Else
9:          Label session as normal
10:         Update C(i) = Legitimate User
11:     End If
12:     Monitor next session activity for continuous authentication
13: End For
14: Return classification results C
```

4. Experimental setup

4.1. Dataset Description

4.1.1. Dataset Overview

The present research uses Insider Threat Dataset of Corporate Environments that is publicly accessible on Kaggle. The dataset was designed in a way that it makes it appear as real-life enterprise user behaviour and insider threat operation. It has organized activity records of interaction of employees with organizational systems, such as events of logins, file access operations, network usage and communication behaviour. Such behavioural records enable to examine a pattern of activity performed by the user and diagnose unusual behaviours that can be linked to insider threats.

The dataset will be comprised of 45,000 instances of behavioural activities, in which each instance is a user activity session in enterprise network environment. A record has more than one behavioural attribute that defines interaction of a user with the resources of a system. A binary threat label is also available in the dataset, whether the activity is normal behaviour (0) or insider threat activity (1). This labelled format supports both experimental and supervised learning as well as anomaly detection [28].

4.1.2. Data Structure and Features

The data is a user-driven behavioural records of daily enterprise operations. Every instance is associated with a user session or a user activity window and consists of numerical and categorical attributes of behavioural patterns. The possible major categories of features are:

- **Authentication Features:** frequency of logins, session length, abnormal timings of login.
- **File Access Behavior:** accessed files, sensitive file operations, suspicious access operations.
- **System Usage Metrics:** resource usage, interaction with devices.
- **Network and Communication Indicators:** frequency of browsing, behaviour of external access.
- **Threat Label:** binary classification label (Normal or Insider Threat)

These structured attributes facilitate the analysis of behavioural modelling and sequences of time.

4.1.3. Data Splitting Strategy

In order to guarantee high-quality performance evaluation and avoid overfitting, the dataset will be separated into two subsets:

 doi:10.20944/preprints202604.1711.v1

- Training Set (80%) - This is used to train the LSTM model and autoencoder. A LSTM net acquires temporal behaviour patterns based on sequence of user activities and auto encoder is mostly trained on regular user examples to acquire default system behavioural traits and to reconstruct normal system activities.
- Testing Set (20%)- It is only utilized when trained models are being evaluated finally. This dataset has hidden cases, which enable evaluation of classification quality, detection of anomalies, and the overall effectiveness of insider threat detection system suggested.

4.2. Data preprocessing

The process creates enterprise activity logs which enable detection of insider threats. The data cleaning procedures handle removal of both incomplete records and inconsistent data from the raw log data which originates from CERT Insider Threat Dataset. The process of feature normalization establishes uniform value ranges through its implementation. The system divides user activities into distinct time intervals which enable the detection of their behavioural patterns while essential features get transformed into structured vectors that LSTM and Autoencoder models use for analysis.

4.2.1. Data Cleaning

Data preprocessing begins with the cleaning of enterprise activity logs to remove noise, inconsistencies, and incomplete records. The behavioural logs obtained from the CERT Insider Threat Dataset may contain missing values or corrupted entries due to logging errors or system failures. Let the raw dataset be represented as Equation (19),

$$W = \{x_1, x_2, x_3, \dots, x_n\}, \tag{19}$$

where each x_i represents an individual activity record generated by a user in enterprise network. Records containing missing or invalid attributes are removed to obtain a cleaned dataset. The cleaned dataset W_c can be expressed as shown in Equation (20).

$$W_c = W - \{x_i \mid x_i \in W \wedge \text{Missing}(x_i) = 1\} \tag{20}$$

where W_c denotes the cleaned dataset, Wrepresents original dataset, x_i is an activity record, and $\text{Missing}(x_i)$ is a function that indicates whether the record contains missing or corrupted values.

4.2.2. Data Normalization

Enterprise behavioural features often have different numerical ranges, which may affect training stability and performance of machine learning models. To address this issue, feature normalization is applied using the min-max scaling method. The normalized feature value x' is computed as denoted in Equation (21).

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{21}$$

where x represents the original feature value, x' denotes the normalized feature value, $x_{\min}$ is the minimum value of the feature in the dataset, and $x_{\max}$ is the maximum value of the feature.

4.2.3. Temporal Segmentation of User Activities

User activities in enterprise systems occur sequentially over time, making it necessary to analyse temporal behavioural patterns. Let the sequence of user activities be represented as shown in Equation (22).

$$U = \{u_1, u_2, u_3, \dots, u_T\} \tag{22}$$

where U denotes the chronological sequence of user activities and u_t represents the user activity at time step t. To capture temporal patterns, the activity logs are segmented into fixed-size windows. The segmented behavioral sequence S_k can be expressed as depicted in Equation (23).

 doi:10.20944/preprints202604.1711.v1

$$S_k = \{u_k, u_{k+1}, u_{k+2}, \ldots, u_{k+w-1}\} \quad (23)$$

where S_k represents the segmented sequence, k denotes the starting index of the window, and w represents the predefined window size.

4.2.4. Feature Vector Construction

After segmentation, the processed activity logs are converted into structured feature vectors representing behavioural attributes such as login frequency, file access behavior, email interactions, and network usage. The feature representation for each behavioural sequence is defined as shown in Equation (24).

$$F_i = [f_1, f_2, f_3, \ldots, f_m] \quad (24)$$

where F_i denotes the feature vector corresponding to the i^{th} behavioural sequence, f_j represents the j^{th} extracted feature, and m indicates the total number of behavioral features used in the model. The feature vectors serve as inputs for deep learning models which include LSTM and Autoencoder to detect abnormal behavioral patterns that are linked to insider threats.

4.3. Computational Environment

A high-performance computer-based system was provided to build the insider threat detection framework. Python 3.10.19 was used to implement the framework and many other scientific computing libraries, machine learning libraries, and libraries for data visualization (NumPy 1.24.3; Pandas 2.3.3; Scikit-learn 1.3.0; TensorFlow 2.9.1; Imbalanced-learn 0.12.3; Matplotlib 3.10.8; and Seaborn 0.13.2) were used to build the models. The experiments were performed on a system with an Intel Core i9-14900K Processor (3.20GHz), 32GB of RAM (31.7GB usable), and a 64-bit x64-based OS running Windows 11 Home (Version 25H2, OS Build 26200.7840) and Windows Feature Experience Pack 1000.26100.291.0. The combination of hardware and software provided adequate computational resources to effectively train and evaluate the deep learning and machine learning models that comprise the insider threat detection framework.

4.4. Hyperparameter Configuration

The study provides configuration parameters for LSTM-Autoencoder hybrid framework which detects insider threats. The LSTM model uses a sequential architecture which requires an input dimension of (1, features) because each input vector contains behavioural features obtained from user activity logs during a monitoring period. The LSTM layer in this setup contains 64 units which enables the model to learn nonlinear connections between different user activity behavioural patterns. The system uses a dense layer that contains 32 neurons with ReLU activation to identify advanced feature relationships which leads to a sigmoid output layer that classifies normal and malicious behavior. The system uses a dropout rate of 0.3 to stop model overfitting. The Adam optimizer trains model with a learning rate of 0.001 and binary cross-entropy loss for 30 epochs with a batch size of 64. The autoencoder system uses a symmetric encoder-decoder design that follows pattern of 32→16→32 to create compressed representations of standard behavioural patterns. The system conducts training for 30 epochs while utilizing a batch size of 32 and it uses a linear output layer together with mean squared error (MSE) as its evaluation method and Adam optimizer. This system uses the 95th percentile of reconstruction errors to establish an anomaly detection threshold which detects abnormal user behavior. The hybrid detection system uses weighted fusion to combine both model outputs because the LSTM component produces 70% of final score while autoencoder accounts for 30%. This hybrid approach uses supervised behavior classification together with unsupervised anomaly detection to enhance insider threat activity detection.

 doi:10.20944/preprints202604.1711.v1

4.5. Performance Evaluation Metrics

The proposed LSTM-Autoencoder continuous authentication system evaluation uses standard classification performance metrics to measure its effectiveness. The insider threat detection system works as a binary classification system which identifies Normal and Insider Threat situations so it uses multiple complementary metrics to achieve complete system evaluation.

Accuracy measures it represented by Equation (25),

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{25}$$

Accuracy provides a general performance indicator but may be misleading in imbalanced datasets.

Precision measures how many detected insider threats are actually malicious is defined in Equation (26):

$$\text{Precision} = \frac{TP}{TP+FP} \tag{26}$$

High precision delivers two main advantages because it minimizes false positive detection which protects organizations from spending resources on unneeded security alerts.

Recall evaluates how many actual insider threats are correctly detected as shown in Equation (27):

$$\text{Recall} = \frac{TP}{TP+FN} \tag{27}$$

Accordingly, detecting high explosive materials is essential for baggage security.

The F1-score balances precision and recall using their harmonic mean shown by Equation (28),

$$\text{F1-Score} = \frac{2\times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{28}$$

The ROC curve assesses model's discriminant ability over multiple thresholds.

(TPR) shown by Equation (29),

$$\text{TPR} = \frac{TP}{TP+FN} \tag{29}$$

(FPR) illustrated by Equation (30),

$$\text{FPR} = \frac{FP}{FP+TN} \tag{30}$$

The Area Under the Curve (AUC) measures how well two different types of behavior can be separated from each other. A higher AUC value indicates better classification performance.

5. Results and Discussion

5.1. Behavioural Risk and Continuous Authentication Analysis

The behavioural risk assessment tests how well proposed system identifies normal user behavior and detects possible insider threats. The system establishes immediate risk evaluations by tracking user activities across various time intervals which help identify patterns that diverge from established normal patterns. The continuous authentication system needs users to be verified throughout their session because it does not rely on their first authentication check. The dynamic monitoring system identifies insider threats with greater efficiency while maintaining consistent authentication processes during operational activities.

The unified anomaly scoring system divides normal user behavior from insider threat activities according to their risk score distribution. The **Figure 4** shows that normal user activities are concentrated near low-risk values between 0.00 and 0.05, which shows users perform their activities according to standard procedures without any noticeable changes. Insiders commit authentically suspicious activities which define specific threat methods that operate between 0.60 to 0.70 range with 0.65 being their most frequent value. The two distributions show a distinct separation which proves that anomaly detection system successfully identifies normal user behavior and detects unauthorized activities. The detection framework shows successful performance because it correctly

 doi:10.20944/preprints202604.1711.v1

identifies suspicious users through its decision thresholds while maintaining low false alarm rates in enterprise environments.

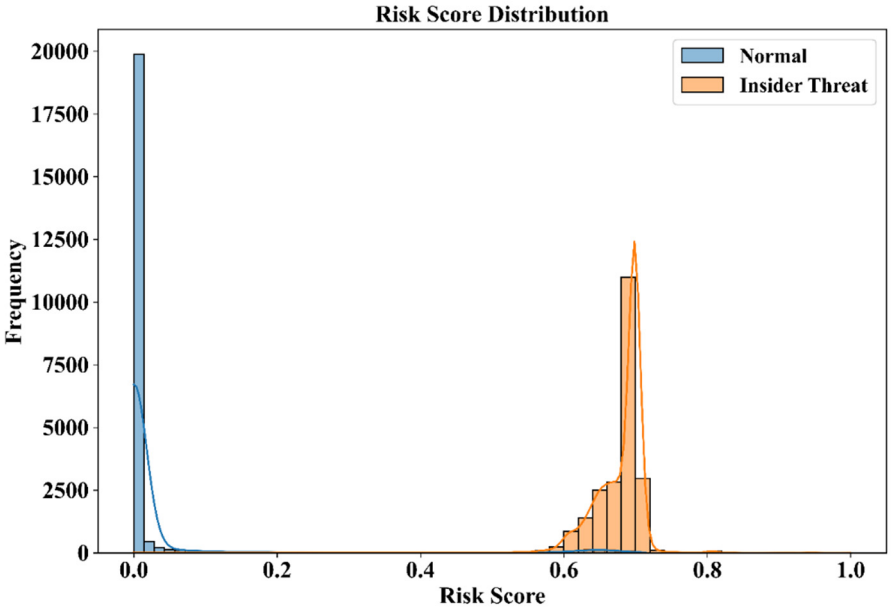


Figure 4. Behavioural Risk Score Distribution for Normal and Malicious Users.

The hybrid LSTM–Autoencoder system produces constant threat detection results because it continuously monitors insider threat risks throughout its entire assessment duration. The raw risk scores in **Figure 5** display a range from 0.60 to 0.70 which shows how different activities were performed during specific timeframes. The aggregation mechanism leads to an established risk score which maintains a range between 0.66 and 0.70 while diminishing sudden changes and unpredictable elements present in the initial predictions. The aggregation process enhances risk assessment accuracy by removing temporary irregularities which would otherwise disrupt essential patterns of behavior. The results demonstrate that proposed model delivers consistent insider threat risk assessments which help organizations improve their ongoing authentication processes throughout their operational environments.

The risk score progression demonstrates how users behave normally through time while showing their possible insider threat activities which are displayed in **Figure 6**. The insider threat risk curve maintains consistently high values between 0.85 and 0.98 which proves the model's capability to detect suspicious conduct patterns that dangerous insiders demonstrate. The normal user risk scores stay at much lower levels which usually fall between 0.00 and 0.15. The proposed hybrid framework shows its ability to differentiate between regular and harmful conduct through its clear ability to separate both types of behavior. The findings show that system successfully monitors behavioural risk trends throughout time to detect insider threats while reducing misidentification.

The authentication stability curve demonstrates system's ability to deliver consistent authentication results across multiple user session durations which **Figure 7** displays. The authentication scores show two peaks between 0.30 and 0.75 because users exhibit different behavioural patterns throughout their sessions. The system detects valid authentication through higher values which show stronger protection, but it experiences temporary drops when users exhibit abnormal patterns of behavior. The continuous authentication system shows stable functioning through normal behavior changes because it successfully detects suspicious activities. The system requires this stability to provide continuous secure access to enterprise systems.

 doi:10.20944/preprints202604.1711.v1

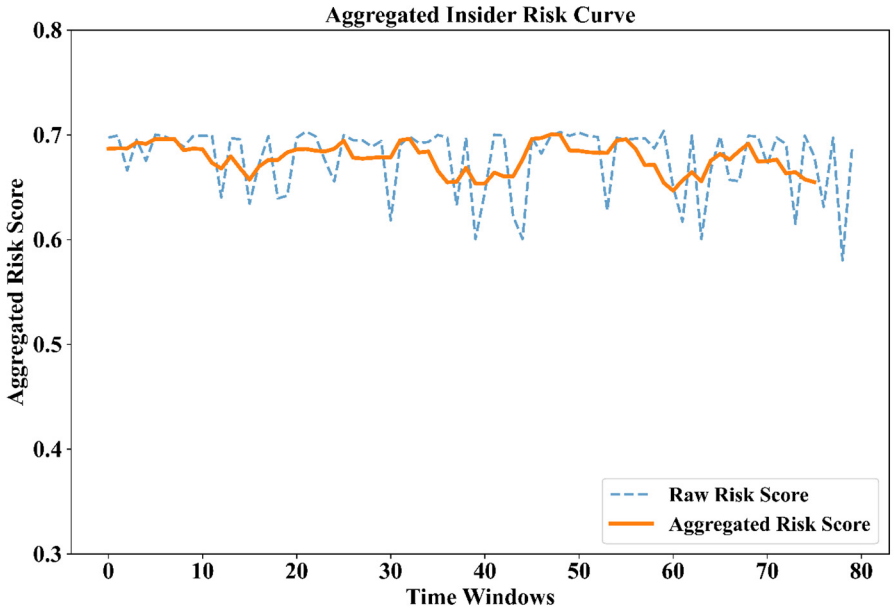


Figure 5. Aggregated Insider Risk Score Analysis.

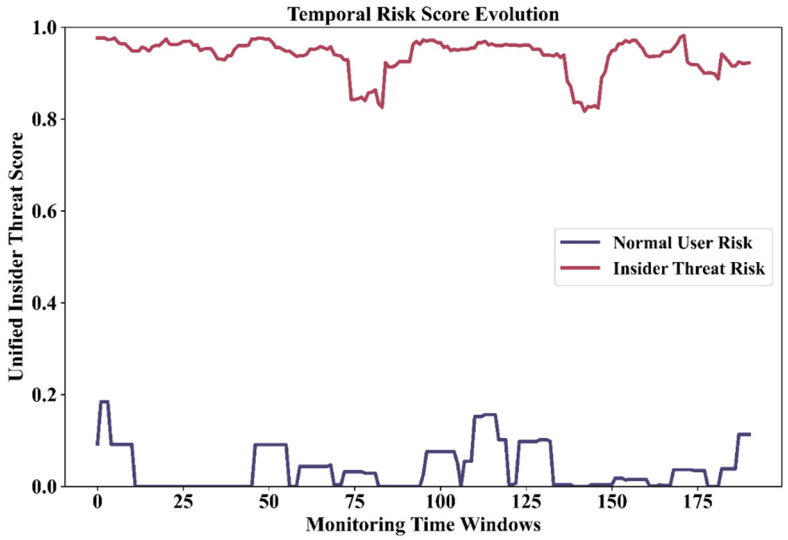


Figure 6. Temporal Evolution of Insider Threat Risk.

The study measures user behavior steadiness through two distinct risk evaluation methods which assess normal users and insider threat activities during extended periods that are shown in **Figure 8**. The insider threat risk trend remains consistently high between 0.84 and 0.98, indicating persistent anomalous behavioural patterns detected by the system. The normal user risk trend maintains a range between 0.00 and 0.15 which shows that users exhibit consistent and predictable patterns of behavior. The two risk patterns show distinct characteristics which prove that proposed model can effectively identify normal user behavior and potential insider threats. The results confirm that the hybrid LSTM–Autoencoder framework successfully captures behavioural stability patterns and enhances the accuracy of continuous authentication systems.

 doi:10.20944/preprints202604.1711.v1

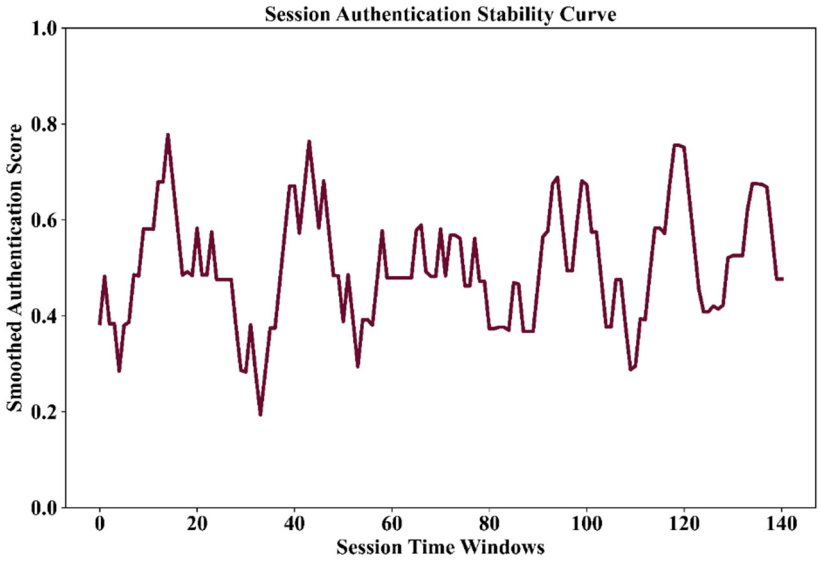


Figure 7. Session Authentication Stability Evaluation.

The false positive trend graph evaluates the reliability of the proposed insider threat detection framework across multiple monitoring windows. **Figure 9** demonstrates that the false positive rate remains low throughout the monitoring period because most measurements stay between 0.02 and 0.08 and only reach 0.20 to 0.23 in extremely rare instances. The overall average false positive rate remains around 0.05 which shows that the model successfully reduces threat alarms that are incorrectly detected as real threats. The security systems of enterprises must maintain low false positive rates because excessive security alerts will reduce user trust in the detection system and generate additional tasks for security personnel. The observed trend confirms that the proposed hybrid model achieves stable anomaly detection performance with minimal false alarms.

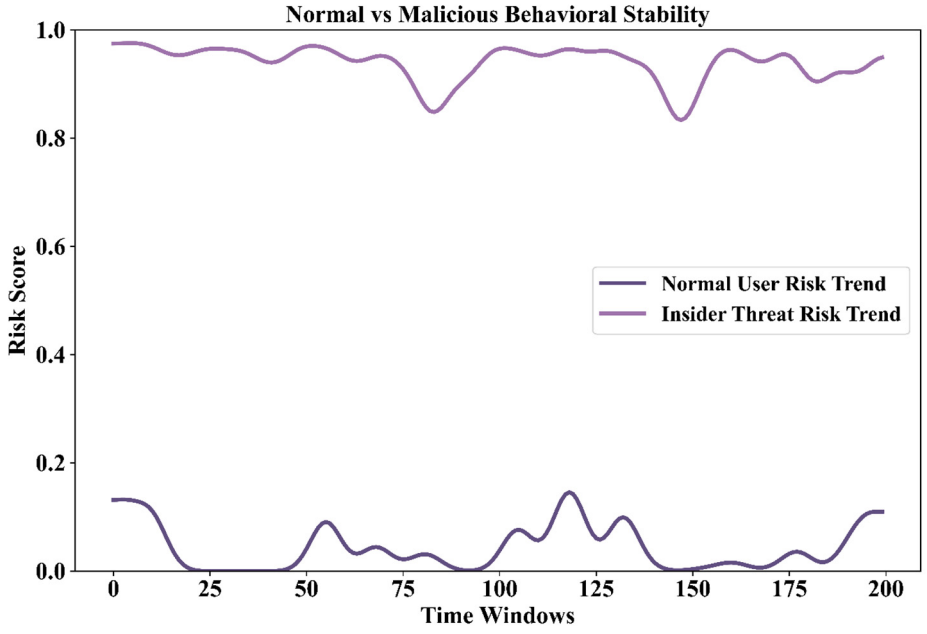


Figure 8. Behavioural Stability Comparison Between Normal and Malicious Users.

 doi:10.20944/preprints202604.1711.v1

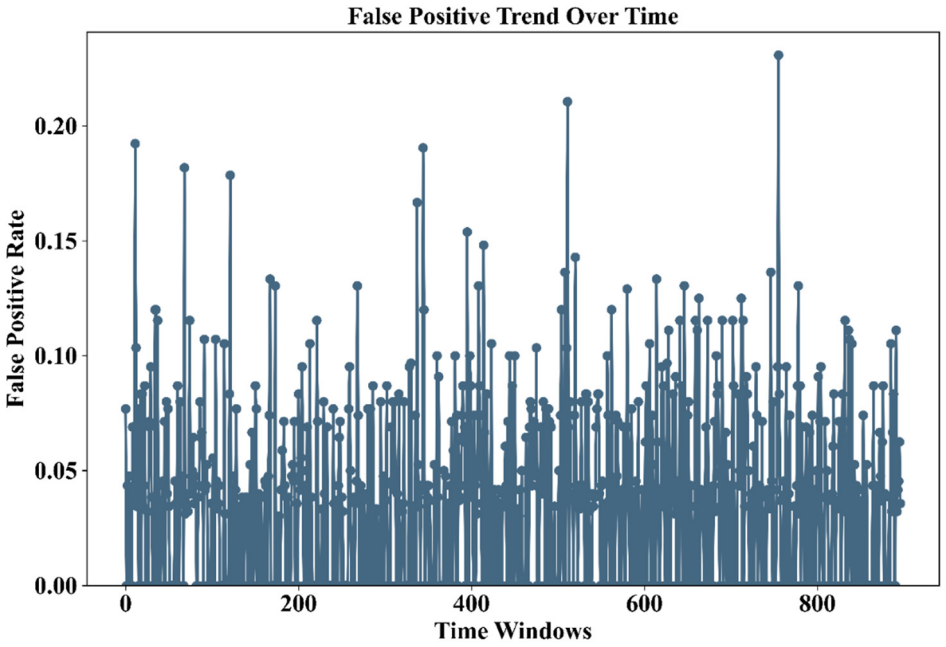


Figure 9. False Positive Trend Analysis Over Time.

5.2. Insider Threat Detection Performance Evaluation

The assessment of performance evaluates how well the proposed framework detects insider threats in enterprise systems. The experimental results demonstrate that the system successfully identifies threats with high accuracy because it achieved strong performance results in accuracy and precision and recall and F1-score measurements throughout the evaluation process.

The performance metrics provide complete assessment of the insider threat detection framework which **Figure 10** displays. The system achieved an accuracy of 97.65% which shows that most user behavior instances received correct classification. A precision score of 96.35% shows that most predicted insider threats correspond to actual malicious activities which reduces unnecessary security alerts. The model demonstrates strong capability to detect insider threats because it achieved a 99.05% recall value. The F1-score of 97.68% shows the system performs equally well in precision and recall while the ROC–AUC value of 99.20% shows a system can effectively distinguish between normal and malicious activities. The results were fully demonstrated by the suggested framework enables enterprises to achieve continuous authentication which protects their security systems.

The classification results of the insider threat detection system of **Figure 11** shows better performance to actual user behavior patterns, which the system tested through its predicted user behavior performance. The model correctly classified 21658 normal user activities while 788 instances were incorrectly identified as insider threats. The system achieved accurate detection of 22218 insider threat activities, which resulted in 228 malicious samples being misclassified as normal behavior. The system developed by the researchers shows better ability to differentiate between normal and harmful behavior patterns. The high accuracy of the LSTM–Autoencoder system comes from its ability to identify complex behavioral patterns that businesses use for their authentication processes. The test results show that LSTM–Autoencoder architecture can successfully identify complex enterprise behavioral patterns while maintaining continuous authentication process trustworthiness.

 doi:10.20944/preprints202604.1711.v1

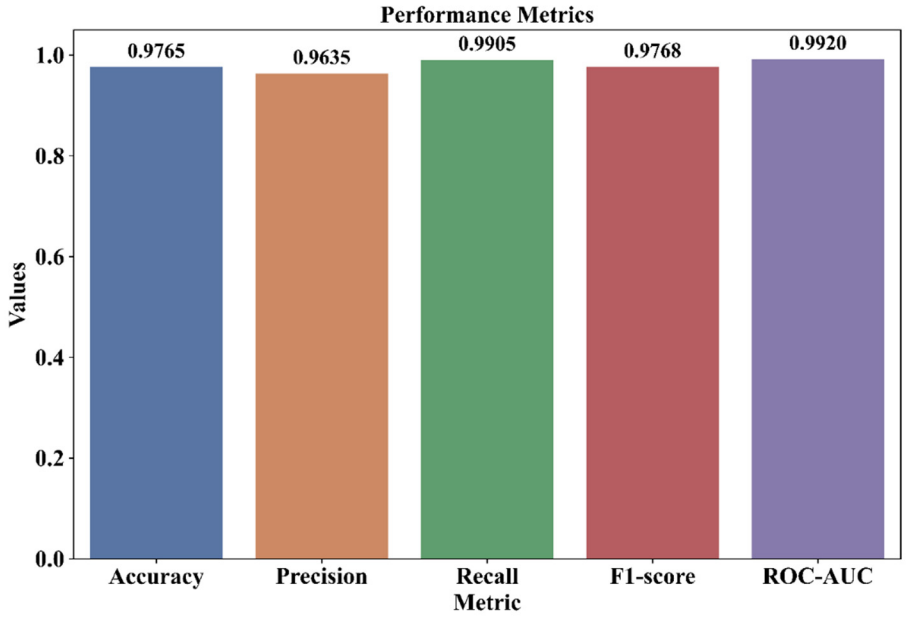


Figure 10. Overall Performance Metrics.

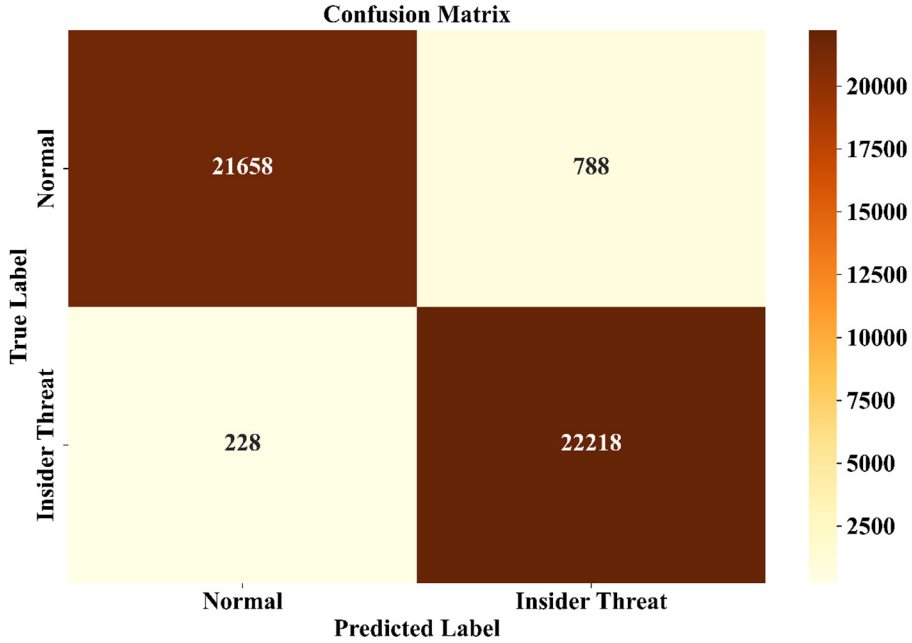


Figure 11. Confusion Matrix of the Proposed LSTM–Autoencoder Framework.

The detection system shows its operational reliability through business environments because the system's false positive rate and false negative rate results which appear in **Figure 12**. The model achieved a false positive rate of approximately 3.75% which showed that the system only falsely identified a small number of actual user activities as suspicious behavior. The system recorded a false negative rate of 0.95% which showed that it successfully detected almost all insider threat activities in its operational environment. The detection system requires a low false negative rate because undetected malicious activities can lead to major security breaches. The results showed that proposed framework achieves its goal of maintaining security protection while enabling system operational stability.

 doi:10.20944/preprints202604.1711.v1

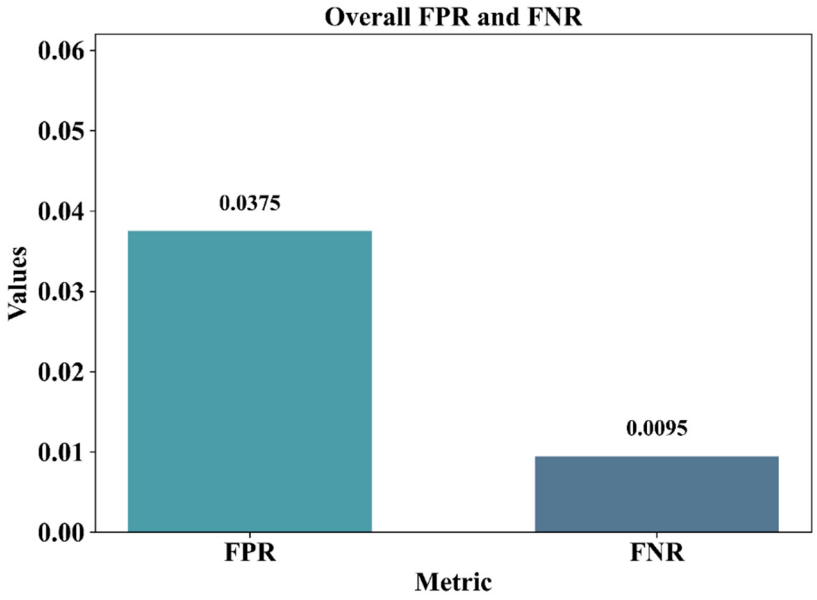


Figure 12. False Positive Rate (FPR) and False Negative Rate (FNR).

The **Figure 13** illustrate learning behavior during the training process. This model achieves its first accuracy improvements when it starts to identify the basic usage patterns of enterprise users. The training process leads to stable training and validation accuracy results which show that the model has reached its final performance level. The model demonstrates strong generalization ability because its behavior matches unknown data patterns and this result proves that the system lacks excessive overfitting. The final validation accuracy approaching approximately 97.6% confirms the robustness of the proposed framework for continuous authentication and insider threat detection tasks.

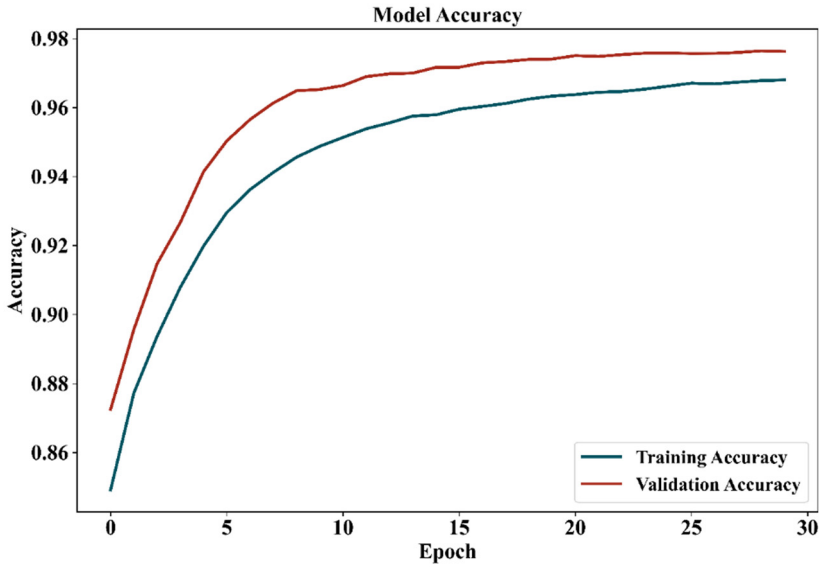


Figure 13. Training and Validation Accuracy Across Epochs.

The **Figure 14** demonstrate how the proposed deep learning model reaches its training goals through testing learning performance. The training process begins at epoch 0 with training loss set to 0.35 and verification loss at 0.29 which shows high error rates until the model acquires dataset behavioral knowledge. The training process shows effective parameter optimization through the

 doi:10.20944/preprints202604.1711.v1

continuous decrease of both loss values. The training loss at epoch 10 reaches 0.14 and the validation loss decreases to 0.11 which shows that prediction accuracy has improved. The training process reaches its final stage at epoch 29 when training loss decreases to 0.10 and validation loss reaches 0.08. The small gap between training and validation loss confirms that the model generalizes well without significant overfitting which proves that the proposed LSTM–Autoencoder framework reaches stable convergence.

The **Figure 15** evaluates the classification performance of different baseline models and the proposed hybrid framework for insider threat detection. The results show that the Logistic Regression model achieves an Average Precision (AP) score of 0.9090, while the Random Forest model improves performance with an AP of 0.9513. Deep learning models demonstrate superior detection capability, with the LSTM model achieving an AP of 0.9898, indicating high precision across a wide recall range. The Autoencoder model achieves an AP score of 0.8092 which shows its independent performance in anomaly detection. The proposed LSTM–Autoencoder hybrid model achieves the highest AP score of 0.9898, maintaining nearly perfect precision for most recall values. The results demonstrate that combining temporal behavioral modeling with reconstruction-based anomaly detection achieves better results for detecting insider threats.

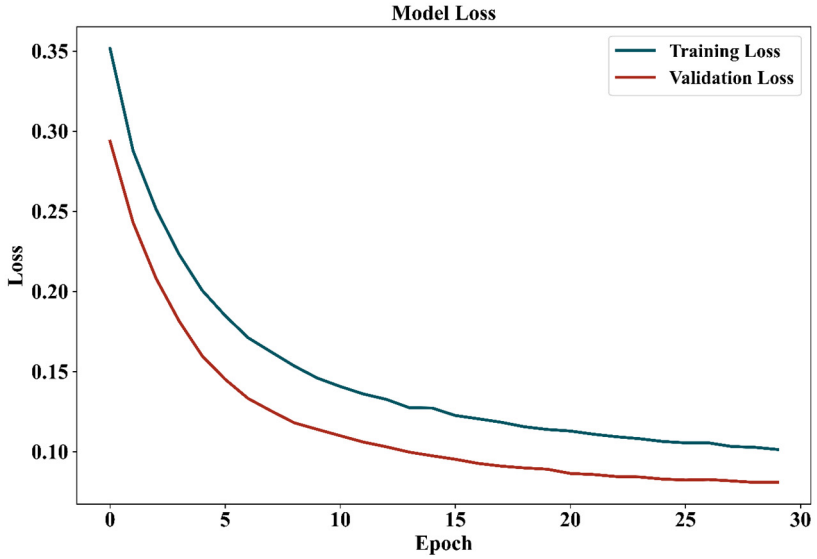


Figure 14. Model Training and Validation Loss Convergence.

The **Figure 16** evaluates the discrimination capability of different classification models by measuring the trade-off between the true positive rate and the false positive rate. These results show that Logistic Regression achieves an AUC value of 0.8898, which demonstrates moderate classification performance. The Random Forest model achieves better results because it has an AUC of 0.9422, which shows that its predictive ability has improved. The Autoencoder model records the lowest AUC value of 0.8076, which demonstrates its inability to independently capture complicated behavioral patterns. The LSTM model achieves a high AUC value of 0.9922, which shows its exceptional ability to track temporal patterns in user behavior sequences. The proposed LSTM–Autoencoder hybrid framework achieves the highest AUC score of 0.9923, which demonstrates its ability to differentiate between normal users and insider threats while proving the continuous authentication system works effectively.

 doi:10.20944/preprints202604.1711.v1

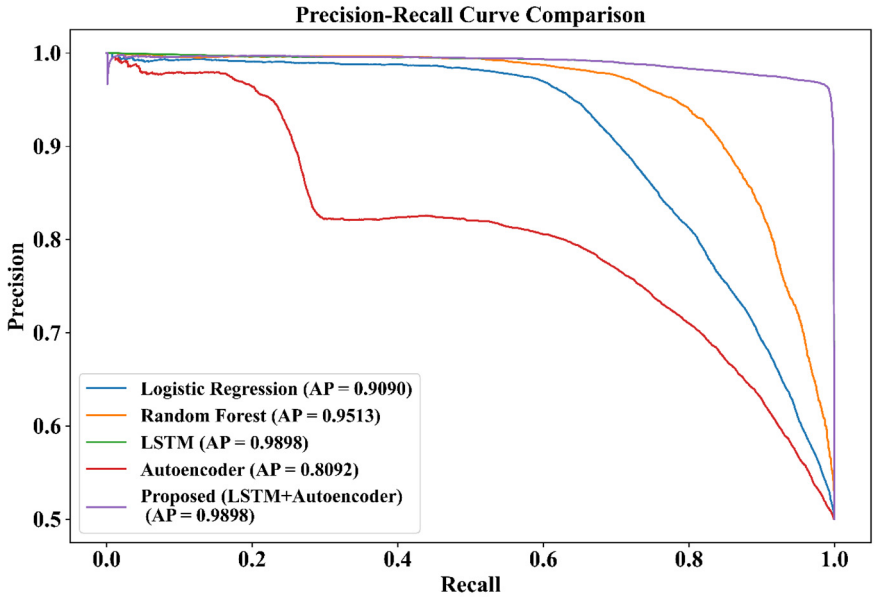


Figure 15. Precision–Recall Curve Analysis of Detection Models.

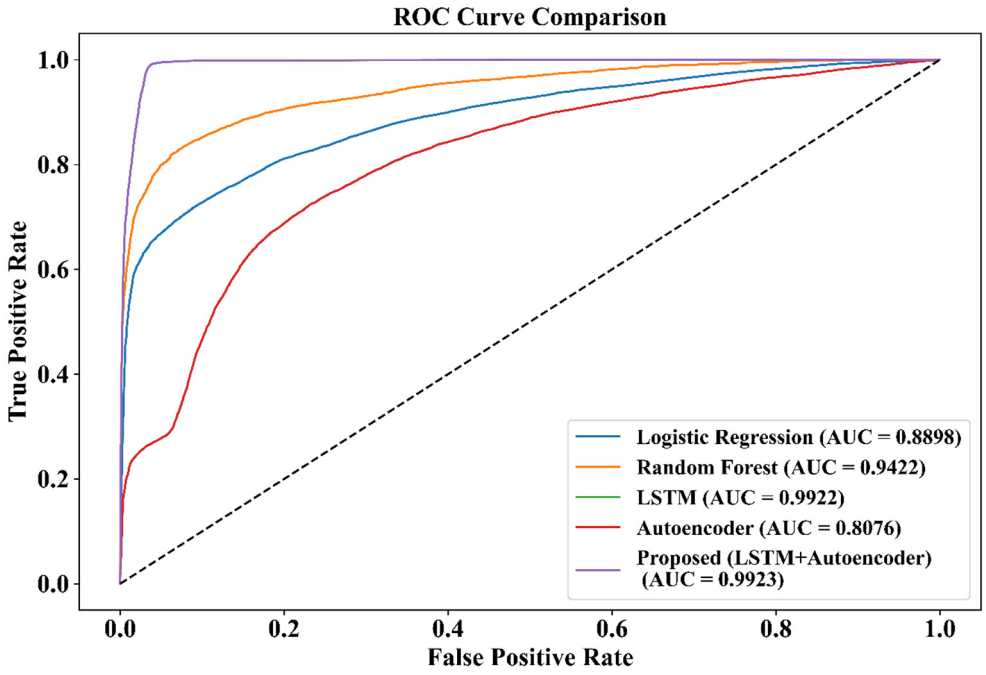


Figure 16. ROC Curve Comparison and AUC Performance.

In this (**Table 1**), we can see how the original hybrid insider threat detection hybrid approach combines baseline machine-learning and deep-learning models to create a new hybrid framework. The traditional logistic regression and random forest results reflect moderately effective performance for both models, as evidenced by the 0.91 and 0.94 accuracies achieved, respectively. In contrast, the deep-learning LSTM model demonstrates superior performance with its ability to learn and recognize temporal behavioural characteristics in an accuracy of 0.975 and an ROC–AUC of 0.99. The standalone autoencoder model exhibits limited performance since it utilizes reconstruction-type anomaly detection exclusively, as such, it lacks the temporal modelling capabilities of the hybrid model, thus limiting its potential effectiveness for insider threats. The hybrid LSTM-autoencoder model combines both temporal-based sequences and behavioural reconstruction elements; therefore,

 doi:10.20944/preprints202604.1711.v1

the overall method achieves the highest levels of accuracy (0.9765), precision (0.9635), recall (0.9903), F1-score (0.9758) and ROC–AUC (0.992). Consequently, this illustrates how the use of combined temporal-based sequence modelling and behavioural reconstruction optimizes expert systems used to detect security threats pertaining to insider actors within an enterprise-based environment.

Table 1. Performance Comparison of Baseline Models and Proposed Insider Threat Detection Framework.

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Logistic Regression	0.91	0.9	0.92	0.91	0.93
Random Forest	0.94	0.95	0.94	0.945	0.96
LSTM	0.975	0.97	0.985	0.977	0.99
Autoencoder	0.89	0.85	0.88	0.865	0.81
Proposed (LSTM + Autoencoder)	**0.9765**	**0.9635**	**0.9905**	**0.9768**	**0.992**

The performance comparison of baseline models Logistic Regression, Random Forest, LSTM, and Autoencoder against the proposed hybrid LSTM–Autoencoder framework is illustrated in **Figure 17**. The traditional models that include Logistic Regression and Random Forest produce moderate results because their accuracy and F1-scores reach between 0.90 and 0.95. The LSTM and Autoencoder models demonstrate better results through their recall and ROC-AUC performance because these models can track time-based patterns while identifying unusual activities. The hybrid framework demonstrates superior performance against all baseline models according to all assessment metrics because it achieves the highest values in accuracy and precision and recall and F1-score and ROC-AUC measurements. The combination of sequential behavioral modeling with reconstruction-based anomaly detection methods demonstrates its capability to enhance insider threat detection through strengthened detection methods. The integrated LSTM–Autoencoder system successfully tracks both temporal user behavior patterns and minor user behavior changes which makes it appropriate for continuous authentication processes and risk management in business setting.

The ablation analysis is meant to point examine the contributions of various components in the proposed insider threat detection framework, as provided in **Table 2**. The baseline model using behavioural features with a machine learning classifier achieves moderate performance, while the LSTM model significantly improves results by capturing sequential behavioural patterns. The standalone autoencoder shows lower performance because it depends on reconstruction-based anomaly detection for its operation. The combination of LSTM and Autoencoder with continuous authentication scoring enhances detection capabilities. The proposed full framework achieves the best performance with an accuracy of 0.9765 and ROC–AUC of 0.992, demonstrating the effectiveness of integrating temporal modelling, anomaly detection, and continuous behavioural risk assessment.

Table 2. Ablation Study of the Proposed Insider Threat Detection Framework.

Model Configuration	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Behavioral Features + ML Classifier	0.912	0.901	0.918	0.909	0.926
LSTM Only (Sequential Behavior)	0.975	0.97	0.985	0.977	0.99
Autoencoder Only (Anomaly Detection)	0.89	0.852	0.88	0.865	0.812
LSTM + Autoencoder (No Risk Fusion)	0.972	0.961	0.982	0.971	0.988
+Continuous Authentication Scoring	0.975	0.963	0.987	0.975	0.991
Proposed Full Framework	**0.9765**	**0.9635**	**0.9905**	**0.9768**	**0.992**

 doi:10.20944/preprints202604.1711.v1

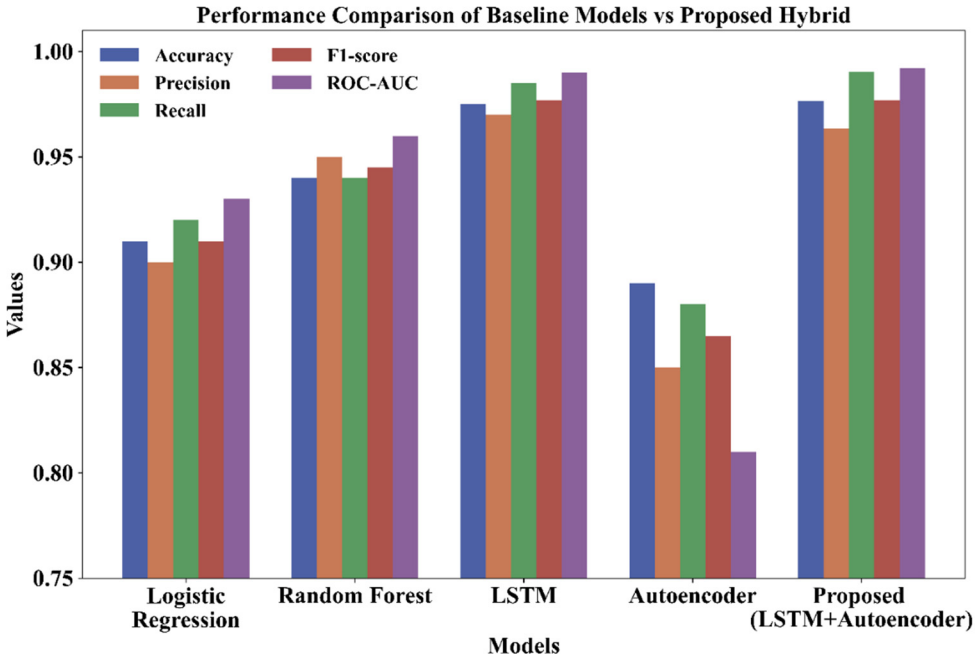


Figure 17. Comparison of Baseline Models Proposed Hybrid Framework.

5.3. Discussion

The proposed behavioural biometrics–based insider threat detection framework effectively detects abnormal user activities in enterprise environments according to the experimental results. The system uses LSTM-based sequential behaviour modelling together with autoencoder-based anomaly detection to detect both temporal behaviour patterns and abnormal user activity. The results indicate that deep learning approaches outperform traditional machine learning models in detecting complex insider threat behaviours. Continuous authentication combined with dynamic risk scoring enhances detection accuracy through the assessment of behavioural variations throughout different time periods. The ablation finding results demonstrate that all framework components boost system performance while the complete model demonstrates superior accuracy and ROC-AUC results. The window-based behavioural analysis system in the proposed framework enables incremental processing of user activity logs despite using an offline evaluation environment for experiments. The model supports expansion of enterprise conditions with extremely sizable user activity data emanating continuously. Enterprise monitoring systems which include log management and SIEM platforms can use deep learning models after their training because the inference phase demands minimal computational resources. The research findings show that the proposed framework serves as an effective security solution which scales to handle continuous insider threat monitoring in enterprise systems.

6. Conclusion and Future Works

The primary objective of this research is a framework that incorporates behavioral characteristics for detecting insider threats continuous authentication and anomaly detection utilizing an LSTM-Autoencoder system. The LAE system analyses user behaviors such as authentication activities, access to files, and other session events to create real-time assessments of insider threats, thus helping to determine whether there has been an occurrence of a security breach as opposed to normal user behavior. The LAE modelling was tested using a corporate environment insider threat data set and proved to be a very useful predictor of normal and abnormal firewall activity with an overall accuracy of 97.65%, a precision of 96.35%, recall of 99.05%, and F1 score of 97.68%. The LAE modelling was

 doi:10.20944/preprints202604.1711.v1

found to have a ROC-AUC value of 99.20%, indicating a high degree of ability to discriminate between normal activities and malicious actions performed by personnel within an organization, while also presenting a low false positive rate. Furthermore, the LAE modelling tracked user behaviors over time, and provided evidence of the existence of potential insider threats through providing a measure of the ongoing stability of session authentication and behavioural risk score analysis. As such, this research demonstrates that LSTM-Autoencoder framework can enable businesses to create a secure and scalable solution for continuous detection of insider threats within operational environment of a company and to provide the necessary resources to mitigate an impact of those insider threats on a company's financial resources.

The proposed insider threat detection framework has strong detection capabilities. However, it requires additional enhancements in order to perform effectively in real-world contexts. Future research will incorporate more contextual behaviour data into the creation of risk score estimates and will also incorporate more advanced deep learning models into the solution in order to improve the ability to track extended behavioural patterns of individuals. The proposed system will also be tested within actual business settings where the system will allow for ongoing surveillance/monitoring of users as well as dynamic assessment of risk. Future research will also explore the use of federated learning methods, which allow multiple organisations to work together on the detection of insider threats while maintaining confidentiality of their data. As a result of these improvements, the developed insider threat detection system will also be scalable, reliable, and usable in 'in-use' environments.

Data Availability Statement: https://www.kaggle.com/datasets/ahmeduzaki/insider-threat-dataset-for-corporate-environments

References

1. P. Manoharan, W. Hong, J. Yin, H. Wang, Y. Zhang, and W. Ye, "Optimising Insider Threat Prediction: Exploring BiLSTM Networks and Sequential Features," *Data Sci. Eng.*, vol. 9, no. 4, pp. 393–408, Dec. 2024, doi: 10.1007/s41019-024-00260-z.
2. W. Hong et al., "A graph empowered insider threat detection framework based on daily activities," *ISA Trans.*, vol. 141, pp. 84–92, Oct. 2023, doi: 10.1016/j.isatra.2023.06.030.
3. L. Yuan et al., "FusionITD: enhanced cross-modal insider threat perception framework via behavior-semantic fusion," *Cybersecurity*, vol. 9, no. 1, p. 119, Feb. 2026, doi: 10.1186/s42400-026-00555-w.
4. F. Bamashmoos, "Adaptive Privacy-Preserving Insider Threat Detection Using Generative Sequence Models," *Future Internet*, vol. 18, no. 1, p. 11, Jan. 2026, doi: 10.3390/fi18010011.
5. M. Villarreal-Vasquez, G. Modelo-Howard, S. Dube, and B. Bhargava, "Hunting for Insider Threats Using LSTM-Based Anomaly Detection," *IEEE Trans. Dependable Secure Comput.*, vol. 20, no. 1, pp. 451–462, Jan. 2023, doi: 10.1109/TDSC.2021.3135639.
6. N. Khan, R. J. Houghton, and S. Sharples, "Understanding factors that influence unintentional insider threat: a framework to counteract unintentional risks," *Cogn. Technol. Work*, vol. 24, no. 3, pp. 393–421, Aug. 2022, doi: 10.1007/s10111-021-00690-z.
7. A. Georgiadou, S. Mouzakitis, and D. Askounis, "Detecting Insider Threat via a Cyber-Security Culture Framework," *J. Comput. Inf. Syst.*, vol. 62, no. 4, pp. 706–716, Jul. 2022, doi: 10.1080/08874417.2021.1903367.
8. H. M. Zangana, Z. B. Sallow, and M. Omar, "The Human Factor in Cybersecurity: Addressing the Risks of Insider Threats," *J. Ilm. Comput. Sci.*, vol. 3, no. 2, pp. 76–85, Jan. 2025, doi: 10.58602/jics.v3i2.37.
9. P. Bansal and A. Ouda, "Continuous Authentication in the Digital Age: An Analysis of Reinforcement Learning and Behavioral Biometrics," *Computers*, vol. 13, no. 4, p. 103, Apr. 2024, doi: 10.3390/computers13040103.
10. T. Tian, C. Zhang, B. Jiang, H. Feng, and Z. Lu, "Insider threat detection for specific threat scenarios," *Cybersecurity*, vol. 8, no. 1, p. 17, Mar. 2025, doi: 10.1186/s42400-024-00321-w.

 doi:10.20944/preprints202604.1711.v1

11. K. Randive, R. Mohan, and A. M. Sivakrishna, "An efficient pattern-based approach for insider threat classification using the image-based feature representation," *J. Inf. Secur. Appl.*, vol. 73, p. 103434, Mar. 2023, doi: 10.1016/j.jisa.2023.103434.
12. J. Muzaffar and N. Mazher, "AI-Powered Behavioral Analysis for Insider Threat Detection in Enterprise Networks," *Balt. J. Multidiscip. Res.*, vol. 1, no. 2, pp. 1–11, Jun. 2024.
13. A. Mahfouz, A. Hamdy, M. A. Eldin, and T. M. Mahmoud, "B2auth: A contextual fine-grained behavioral biometric authentication framework for real-world deployment," *Pervasive Mob. Comput.*, vol. 99, p. 101888, Apr. 2024, doi: 10.1016/j.pmcj.2024.101888.
14. M. S. Mohamed and A. Arabo, "A SIEM-Integrated Cybersecurity Prototype for Insider Threat Anomaly Detection Using Enterprise Logs and Behavioural Biometrics," *Electronics*, vol. 15, no. 1, p. 248, Jan. 2026, doi: 10.3390/electronics15010248.
15. S. S. P. Pennada, S. K. Nayak, and V. K. M, "Insider Threat Detection Using Behavioural Analysis through Machine Learning and Deep Learning Techniques," *Int. Res. J. Multidiscip. Technovation*, vol. 7, no. 2, pp. 74–86, Mar. 2025, doi: 10.54392/irjmt2527.
16. S. S. Abba, O. A. Obioha-Val, V. O. Ejiofor, O. M. Olaniyi, and N. R. Mayeke, "Behavioral Biometrics-Powered Continuous Authentication for Zero-trust Remote Work Environments: A Multi-factor Identity Verification Framework," *Asian J. Res. Comput. Sci.*, vol. 18, no. 12, pp. 20–41, Nov. 2025, doi: 10.9734/ajrcos/2025/v18i12788.
17. O. O. Aramide, "AI-Driven Identity Verification and Authentication in Networks: Enhancing Accuracy, Speed, and Security through Biometrics and Behavioral Analytics," *ADHYAYAN J. Manag. Sci.*, vol. 13, no. 02, pp. 60–69, Aug. 2023, doi: 10.21567/adhyayan.v13i2.10.
18. I. Ibraheem, A. T. Morufat, S. K. Segbefia, and A. A. Abdulrasaq, "Detecting Malicious Insider Threats through Anomaly-Based User Behaviour Analytics in Enterprise Networks: Machine Learning Approach," *South. Afr. J. Secur.*, vol. 3, p. 23 pages-23 pages, Dec. 2025, doi: 10.25159/3005-4222/18099.
19. J. Hu, W. Wu, T. Chuan, and Q. Peng, "Enterprise Internal Threat Authentication Traceability Technology Based on Key Authentication System," *J. Cyber Secur. Mobil.*, vol. 14, no. 3, pp. 623–652, May 2025, doi: 10.13052/jcsm2245-1439.1435.
20. D. He, X. Lv, X. Xu, S. Chan, and K.-K. R. Choo, "Double-Layer Detection of Internal Threat in Enterprise Systems Based on Deep Learning," *IEEE Trans. Inf. Forensics Secur.*, vol. 19, pp. 4741–4751, 2024, doi: 10.1109/TIFS.2024.3372771.
21. K. Saminathan, S. T. R. Mulka, S. Damodharan, R. Maheswar, and J. Lorincz, "An Artificial Neural Network Autoencoder for Insider Cyber Security Threat Detection," *Future Internet*, vol. 15, no. 12, p. 373, Dec. 2023, doi: 10.3390/fi15120373.
22. U. Uslu, Ö. D. İncel, and G. I. Alptekin, "Evaluation of Deep Learning Models for Continuous Authentication Using Behavioral Biometrics," *Procedia Comput. Sci.*, vol. 225, pp. 1272–1281, Jan. 2023, doi: 10.1016/j.procs.2023.10.115.
23. A. Orun, E. Orun, and F. Kurugollu, "Cognitive behavioural characteristics identification for remote user authentication for cybersecurity," *J. Parallel Distrib. Comput.*, vol. 202, p. 105102, Aug. 2025, doi: 10.1016/j.jpdc.2025.105102.
24. X. Tao, Y. Yu, L. Fu, J. Liu, and Y. Zhang, "An insider user authentication method based on improved temporal convolutional network," *High-Confid. Comput.*, vol. 3, no. 4, p. 100169, Dec. 2023, doi: 10.1016/j.hcc.2023.100169.
25. N. Ayanbode, O. A. Abieba, N. Chukwurah, O. O. Ajayi, and A. I. Daraojimba, "Human Factors in Fintech Cybersecurity: Addressing Insider Threats and Behavioral Risks," *Int. J. Multidiscip. Res. Growth Eval.*, vol. 5, no. 1, pp. 1350–1356, 2024, doi: 10.54660/.IJMRGE.2024.5.1.1350-1356.
26. A. F. Baig, S. Eskeland, and B. Yang, "Privacy-preserving continuous authentication using behavioral biometrics," *Int. J. Inf. Secur.*, vol. 22, no. 6, pp. 1833–1847, Dec. 2023, doi: 10.1007/s10207-023-00721-y.
27. T. Al-Shehari, M. Al-Razgan, T. Alfakih, R. A. Alsowail, and S. Pandiaraj, "Insider Threat Detection Model Using Anomaly-Based Isolation Forest Algorithm," *IEEE Access*, vol. 11, pp. 118170–118185, 2023, doi: 10.1109/ACCESS.2023.3326750.

 doi:10.20944/preprints202604.1711.v1

28. Ahmed Mohamed Zaki, "Insider Threat Dataset." Accessed: Mar. 09, 2026. [Online]. Available: https://www.kaggle.com/datasets/ahmeduzaki/insider-threat-dataset-for-corporate-environments
29. S.Sattar, S.Khan, M.I.Khan, A.Akhmediyarova, O.Mamyrbayev, D.Kassymova, D.Oralbekova, J.Alimkulova. Anomaly detection in encrypted network traffic using self-supervised learning. Scientific Reports 15, 26585 (2025). doi: 10.1038/s41598-025-08568-0.

Short Biography of Authors (SBA)

Nursultan Kuldeyev, PhD Candidate in Management Information Systems, Junior Researcher in Kazakh National Technical University named after K. I. Satbayev. In July 2025, he received a masters's degree in Management of Information Systems from the Kazakh National Technical University named after K. I. Satbayev.

He is a specialist in the field of artificial intelligence, machine learning, big data analytics, information systems management, and digital platform development. His research focuses on the creation of intelligent information systems, data-driven technologies, and the development of digital ecosystems for tourism and smart services.

Currently, he is the author of 5 scientific publications and 2 patents related to the application of artificial intelligence, data analysis, and modern information technologies.

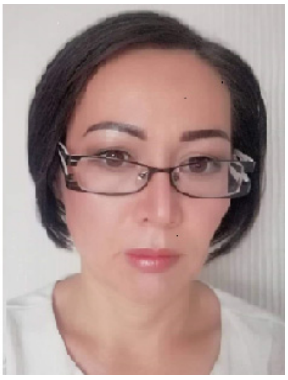
ORCID: https://orcid.org/0009-0004-5906-1040

Orken Mamyrbayev received the B.S. degree and M.S. degree in Information Systems from Abai University, Almaty, Kazakhstan. He obtained his Ph. D. in Information systems at the Kazakh National Technical University named after K. I. Satbayev, and Professor at the Institute of Information and Computational Technologies, Kazakhstan. He has been a Chief Researcher with the Laboratory of computer engineering of intelligent systems at the Institute of Information and Computational Technologies. He is currently a Deputy General Director and head of the Laboratory of computer engineering of intelligent systems at the Institute of Information and Computational Technologies, Kazakhstan. He is a member of the dissertation council "Information Systems" at L.N. Gumilyov Eurasian National University in the specialties Computer Sciences and Information systems. He is the author of five books, more than 130 articles,

 doi:10.20944/preprints202604.1711.v1

and more than 20 inventions and copyright certificates for an intellectual property object in software. His main research field of activity is related to machine learning, deep learning, and speech technologies.

ORCID: https://orcid.org/0000-0001-8318-3794



Ainur Akhmediyarova, PhD, Associate Professor, teaches at Satbayev University. In 2001, she received a master's degree from the Kazakh National Pedagogical University named after Abay, specializing in Mathematics. In July 2018, she received a PhD degree in Information Systems from the Kazakh National Technical University named after K. I. Satbayev.

She is a specialist in the field of mathematical modeling, data processing, information security, pattern recognition, speech recognition, system automation and the development of various software products and software and hardware systems.

Currently, she is the author of more than 120 scientific papers, 3 monographs and 8 copyright certificates.

ORCID: https://orcid.org/0000-0003-4439-7313

Assel Yerzhan, teaches at Satbayev University. In 2018, she received a master's degree from the Samara National Research University named after Academician S. P. Korolev, specializing in Fundamental Informatics and Information Technology.

She is an information security specialist.

ORCID: https://orcid.org/0000-0001-9304-1335