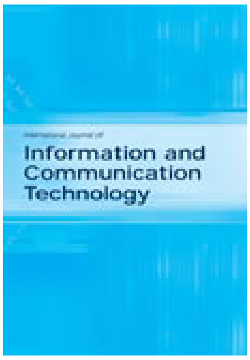


Information and
Communication
Technology

International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Multisource financial data fusion enhanced anomaly transaction detection and early-warning mechanism

Ziwei Rao

DOI: 10.1504/IJICT.2025.10075205

Article History:

Received:	11 August 2025
Last revised:	17 September 2025
Accepted:	19 September 2025
Published online:	12 January 2026

Multisource financial data fusion enhanced anomaly transaction detection and early-warning mechanism

Ziwei Rao

School of Economics and Management,
Wuhan Technical College of Communications,
Wuhan, Hubei, 430000, China
Email: reginarzw@163.com

Abstract: This research proposes a CLST architecture that integrates multiple data sources using Siamese neural networks (SNN) to identify unusual financial transactions. By leveraging spatial, temporal, and multimodal feature learning alongside class imbalance handling, the model outperforms existing methods in recall, F1-score, and precision, enabling a robust early- warning system for fraud prevention. Multisource data fusion enhances detection accuracy by combining complementary information from diverse financial streams. While prior studies have applied rule-based, or machine learning methods to unimodal datasets, and recent multimodal approaches show promise, challenges remain in complex financial networks. The proposed hybrid method combines CNNs, LSTMs, MLPs, and SMOTE to address class imbalance, with SNN-based feature extraction improving robustness. Experiments demonstrate maximum precision of 0.937 and an F1-score of 0.787, with SNN + RF and SNN + SVM outperforming traditional and SMOTE-based models. Statistical analysis confirms SNN-based models achieve superior stability and balanced accuracy in anomaly detection.

Keywords: multisource financial data fusion; anomaly transaction detection; early-warning mechanism; multimodal learning; fraud detection; credit card transactions.

Reference to this paper should be made as follows: Rao, Z. (2025) ‘Multisource financial data fusion enhanced anomaly transaction detection and early-warning mechanism’, *Int. J. Information and Communication Technology*, Vol. 26, No. 49, pp.1–18.

Biographical notes: Ziwei Rao studied in the College of Arts and Sciences of Yangtze University from 2011 to 2015 and received her Bachelor’s in Accounting in 2015. From 2015 to 2017, she studied in Guangdong University of Technology and received her Master’s in Accounting in 2017. She has been working at Wuhan Technical College of Communications since 2017. She has published a total of seven papers. Her research interests are included financial big data analysis.

1 Introduction

1.1 Definition of anomaly detection and its importance in real-world applications

Finding data patterns that are really out of the ordinary is known as anomaly detection (Herden, 2020). Problems with the system, security breaches, fraudulent transactions, or even medical issues can be signalled by these outliers. A prompt and accurate detection of anomalies is critical for preserving system integrity, lowering risks, and assuring operational efficiency in domains such as transportation, healthcare, cybersecurity, and finance.

1.2 Overview of multimodal data

Multimodal data is the result of the growing complexity and diversity of data streams produced by modern systems (Hanchuk and Semerikov, 2025). Structured logs, pictures (like security footage), sounds (like environmental noises or voice commands), data from sensors (like Internet of Things readings), and text are all examples of this. Different imaging modalities pick up on other parts of the system's behaviour; when combined, their distinct insights help fill in the gaps in our knowledge of typical and atypical patterns.

1.3 Why traditional anomaly detection methods fall short with complex, multisource data

Most anomaly detection methods have been developed for use with homogeneous or single-modal data sets (Rella, 2022). The intricate interdependencies, high dimensionality, and variability of multimodal datasets are too much for these approaches to manage. Reduced detection accuracy and missed anomalies are typical results of their inability to capture contextual cues that extend across several senses.

1.4 Purpose and significance of multimodal anomaly detection

In order to overcome these shortcomings, multimodal anomaly detection integrates data from various sources, enabling more thorough contextual analysis and improved anomaly detection accuracy (Chidibere, 2024). Not only does this method make detection more accurate, but it also makes it more resilient when dealing with inadequate or noisy data. It is feasible to find minor or concealed abnormalities that unimodal techniques would overlook by utilising complementary modalities. Data extraction for anomaly detection is one application of automated analysis. A pattern might be considered anomalous when it arises in a sample that differs from the norm or the most common sample (Mikuni, 2024). Inconsistent results, outliers, or anomalies describe these strange patterns. The statistical community initially recognised the need for anomaly detection in the early 1800s, and since then, numerous methods for its detection have been created. Back then, anomaly identification had to be done by hand by experts in each discipline via eye inspection. But there were issues with manual detection as well. The drawbacks of manual detection include, but are not limited to, the following: uncertainty, long detection times, human mistakes, etc. Anomaly detection solutions that use machine learning techniques have recently been developed as a result of those above (Shiva, 2024).

Recently, anomaly detection has emerged as a significant obstacle for deep learning and machine learning. Modern landscapes rely heavily on automated anomaly detection systems because human categorisation becomes impractical due to the sheer volume of samples. Management, astronomical data, visible light curves, credit card fraud, cybersecurity breaches, critical safety system defects, healthcare, insurance, military surveillance, and many more applications are among the many that make use of anomaly detection (Palakurti, 2024). This essay will take a close look at recent research on anomaly detection. Use RNNs to identify outliers in production system time series data. With this technology, makers could identify any irregularities that occurred while the system was running. Using time-series data, the model was able to detect three common types of irregularities in a diesel engine assembly process, which allowed for an evaluation of its performance. In order to find outliers in univariate time series, they suggested a split framework (Manafi, 2025). In this inquiry, time series forecasting was the initial stage. The second stage centred on identifying outliers. To make predictions, the study employed CNNs and LSTM networks (long short-term memory with bidirectional functionality). After detection, the mean absolute error approach was used consistently.

An RLAD (reinforcement learning from pixels for autonomous driving) hybrid deep learning approach was used for anomaly identification. The security of industrial control system (ICS) networks has long piqued the curiosity of academics around the globe, and anomaly detection systems have recently grown in importance. The accuracy of assessing the health state of industrial equipment has been significantly improved in recent years by using anomaly detection approaches that rely on multi-physical quantity fusion (Riegler, 2021). Systematic investigations on communication protocol security and fuzzy testing frameworks have illuminated new approaches to evaluating vulnerabilities in ICS networks. Traditional anomaly detection approaches using offline processing or static analysis are inadequate for meeting the needs of industrial control systems for real-time monitoring and rapid response. There are usually extensive temporal relationships in ICS network traffic data (Koay, 2023). Anomalies in traffic patterns can be effectively detected by leveraging these relationships within time series. The requirement for ICS network anomaly detection in real- time, along with these features, has led to the emergence of real-time traffic prediction as an exciting area of study. This method offers strong support for anomaly identification by enabling continuous monitoring and prediction of ICS network traffic patterns. There are a number of obstacles that this line of inquiry must overcome, though. Traffic data in ICS networks is heterogeneous, nonlinear, and extremely noisy due to the many ways in which industrial control devices operate (Lee, 2023).

Complicating real-time traffic analysis is the fact that data from different devices frequently varies significantly in terms of size, frequency, and sampling methodologies. Data from industrial control systems (ICS) traffic is also very dynamic, changing as production tasks do. Due to these changes, traffic prediction-based anomaly detection algorithms are becoming more and more challenging (Xu and Shang, 2025). Finding abnormalities and correctly retrieving relevant data in real time across such a diverse and ever-changing ICS environment is the main problem. Information about traffic prediction-based real-time anomaly detection algorithms is lacking in the context of industrial control networks. More so, the existing corpus of research is beset by the following issues. By utilising multisource financial data fusion, this research presents a new method for abnormal transaction detection that overcomes the drawbacks of conventional single-modality approaches. The significance of identifying irregularities, such as fraud or system breakdowns, in real-time financial settings characterised by diverse, multi-dimensional, and dynamically produced data is emphasised. An early-warning mechanism that incorporates multimodal data sources to strengthen anomaly detection systems and improve their accuracy and robustness is designed as part of the study's contribution. Using cutting-edge techniques like deep learning models (e.g., LSTM, CNN, VAE), the study demonstrates how combining various data sources improves contextual understanding and decreases false positives.

The research also shows how these methodologies can be applied to financial and industrial control systems, demonstrating how the suggested solution can adapt to complex and dynamic operational situations. Financial fraud prevention, system stability, and strategic decision-making are all greatly enhanced by this fusion-based detection approach, which also provides timely insights. The following structure is used throughout this article: In Section 2, we take a look at what is known about combining financial data from several sources to spot questionable activities. The methodology of the suggested early-warning mechanism is described in Section 3. Section 5 brings the investigation to a close, while Section 4 summarises the results and discusses their ramifications.

2 Literature review

Several theoretical models provide potential definitions of creative accounting. From an accounting standpoint, specific research has shed light on this phrase, highlighting how it represents different approaches to reconciling presentational financial outcomes with the underlying activities (Sabău, 2021). A defining characteristic of creative accounting is the wilful deviation from generally accepted accounting principles in order to achieve a result in reporting. The more you look into it, the more you'll see that these kinds of things happen when businesses attempt to alter their accounting methods from the legal framework to suit their managerial objectives better. Some have proposed a two-tiered understanding, with the first tier addressing efforts to regulate emerging economic phenomena that are not yet accounted for by established accounting rules (Durana, 2022). Generally speaking, this word is defined at the second level as actions that cause financial statement falsification. Innovative Financial Accounting: Its Origins and Applications offers a scholarly perspective. From this point of view, creative accounting is all about getting financial data from its raw, recorded form and making it fit the owners' intended picture. Manipulating legally allowed rules or, in some cases, ignoring specific restrictions, can accomplish this (Urdaneta-Camacho and Guevara-Pérez, 2022).

While innovative bookkeeping practices may help companies manipulate their financial outcomes, this does not necessarily result in monetary benefits, according to another critical assessment that has been added to the discussion. However, such tactics could have a detrimental effect on the company's performance and sustainability in the long run (Blazek and Duricova, 2025). There is a wide range of reasons for creative accounting. Previous studies have shown that financial professionals have systematic behavioural patterns, which supports a shared explanation. These patterns include income smoothing and meeting defined performance targets, in addition to more traditional motivations like tax minimisation and manipulating investor impressions.

2.1 Related work

The stability and dependability of contemporary distributed systems depend on log-based anomaly detection. The principal source of operational intelligence in production environments is system logs (Guo, 2021). This is because microservice designs in these settings can involve hundreds of interconnected components. I can't stress enough how important they are: To start, logs show how the system is behaving in real time, which helps find problems before they affect the quality of service. Secondly, they are crucial for comprehending and avoiding the propagation of faults because they record the intricate interactions among dispersed components. Thirdly, logs are essential for system maintenance and root cause investigation in large-scale deployments since they are the sole complete source of diagnostic information (He, 2021). Using effective log-based anomaly detection, recent industrial studies found that system downtime can be reduced by up to 70% and MTTR by 45%. In cloud computing environments, it is crucial to avoid system failures and maintain service level agreements (SLAs). This is because a single failure could impact several services.

2.2 Standard approaches of identifying abnormalities

Many production settings rely on traditional methods of log-based anomaly detection as their basis for system reliability engineering. Preventing system breakdowns and preserving operational stability have been achieved through the use of these strategies (Xie, 2020). In systems that are vital to the nation's infrastructure, where the ability to spot abnormalities in real time is essential for avoiding catastrophic failures, their significance becomes even more apparent. For instance, by identifying early warning signals in component interactions, classical log analysis methods have effectively avoided system-wide disruptions in large-scale cloud systems. Research conducted by prominent cloud providers indicates that, when executed correctly, proactive anomaly detection using log analysis can avert as many as 85% of possible system failures. Statistical, rule-based, and machine learning techniques are the main categories into which these more conventional approaches fall. In order to spot outliers, rule-based methods specify patterns or thresholds. Although simple, these solutions necessitate a great deal of expertise in the relevant topic and regular manual upkeep (Liu, 2024).

2.3 Traditional methods for anomaly detection

This section will go over the four primary categories of conventional anomaly detection techniques. Methods based on density estimation. One well-known method for detecting outliers is the local outlier factor (LOF) methodology. To deal with complex, multi-dimensional data, the clustering with outlier factor (COF) approach uses the connection principle. The DAGMM algorithm integrates neural networks with a Gaussian mixture model (GMM) to identify anomalies (Xu and Wu, 2021).

Approaches centred on reconstruction. These techniques compare the original data with the rebuilt data in order to identify outliers, after training a model with normal data. Present a hybrid model that combines LSTM and VAE to extract and rebuild features from raw temporal data. Although they utilised a GRU to extract latent features rather than a VAE model, our process was identical to theirs (Wu and Xu, 2021). Anomaly detection is handled by TimesNet using standard algorithms in computer vision, and time-series data is translated from one dimension to two dimensions using a quick Fourier transformation. In addition to traditional methods, generative adversarial networks (GANs) have found usage in fraud detection.

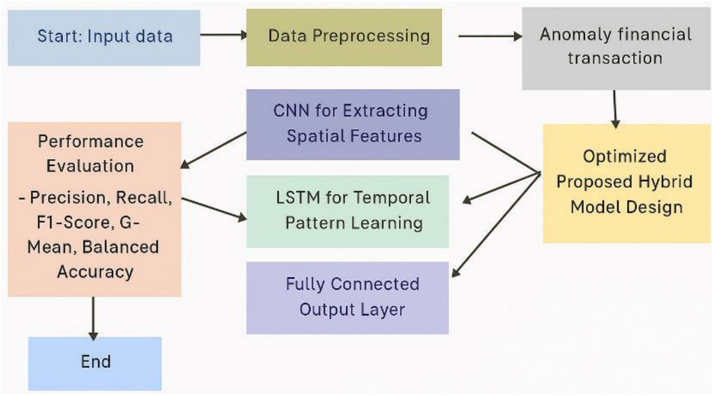
2.4 Multimodal methods for anomaly detection

Recently, multimodal learning (MML) has emerged as a significant field of study due to the advancements in cross-domain data fusion (Zhou and Ma, 2022). Its goal is to strengthen and enhance machine learning and AI systems' capabilities across a range of tasks by making use of complementary information across diverse modalities. Feature vectors from several modalities are typically combined during training by traditional deep learning methods. Each modality is normally processed independently. The feature-level fusion approach disregards the interdependence and complementary nature of modalities due to its oversimplification. One well- known multimodal method for detecting anomalies in industrial settings makes use of 3D point clouds and RGB images (Qu and Liu, 2024). Introduced a novel hybrid fusion approach for multimodal anomaly detection; this scheme fuses RGB features with point cloud features simultaneously, utilising a contrast loss-based unsupervised feature fusion module. When it comes to RGB and visible-light multimodal industrial anomaly detection, given the challenges of representing and dissecting the essential parts unique to each modality while also taking into consideration features that can be shared across modes, a multimodal picture fusion approach was proposed. Developed a multimodal multi-label recognition transformer by integrating a convolutional neural network (CNN) with a transformer; this model can identify numerous things in a single image at the same time.

3 Proposed methodology

We have included a flowchart in this paragraph to help clarify the suggested process. Begin by preparing the data. Then, use a Siamese neural network (SNN) to extract features. Next, classify the data using several models. Finally, evaluate the findings. Figure 1 is a flow diagram that shows the whole procedure.

Figure 1 An illustration of the process that is suggested for detecting anomalies (see online version for colours)


Start: Input data
Data Preprocessing
Anomaly financial transaction
CNN for Extracting Spatial Features
Performance Evaluation
- Precision, Recall, F1-Score, G-Mean, Balanced Accuracy
LSTM for Temporal Pattern Learning
Optimized Proposed Hybrid Model Design
Fully Connected Output Layer
End

3.1 Data preprocessing

Preprocessing is essential for successful modelling due to the high class imbalance and anonymised data:

- *Normalisation:* preprocessing is essential for successful modelling due to the high-class imbalance and anonymised data:

  $$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

  Reshaping: For CNN compatibility, the data was transformed into 2D arrays. Every instance followed a 6 × 5 matrix structure.

- *Class imbalance handling:* synthetic minority over-sampling technique (SMOTE) (Chou, 2025) used in conjunction with undersampling allows a more equitable training set to be generated to tackle the significant class disparity.

- *Train-test split:* to ensure stratified sampling to maintain the class ratio, the data was split 80/20 between the training and test sets.

  $$y_i\left(w.x_i + b\right) \geq 1 - \xi_i, \xi_i \geq 0, \tag{2}$$

3.2 Anomaly financial transaction

Financial transactions across time can be effectively modelled using transaction networks, which are typically depicted as weighted directed temporal networks. In this type of network, a path is an ordered set of *n* separate edges that connects two nodes i and j across a time interval T. Graphically, this kind of network might look like GT = (V, LT).

$$\begin{gathered} P_{i,j}^{T} = \left\{\left(i, v_i\right), e\left(v_1, v_2\right), e\left(v_2, v_3, \cdots, e\left(v_{n-1,j}\right)\right)\right\} \\ \text{with } v_x \in V \text{ and } e\left(v_x, v_{x+1}\right) \in L_T \text{ for all } x \in \{1, \cdots, n-1\} \end{gathered} \tag{3}$$

Next, we find the path's weight, $P_{i,j}^{T}$ by

$$\sum_{i}^{n-1}\left(e\left(v_i, v_{i+1}\right)\right) \tag{4}$$

To represent the weight of each edge along the path P_i, it has $W(P_i) = \{w(e(v_1, v_2)), \ldots, w(e(v_{n-1}, v_n))\}$. Every element *i* and every element *j* in graph *G* can be found in the set $Path_{ij}^{(n)}\left(G_T\right)$ if there is a pathway in *G* of length *n*. In actual situations, agents often employ intermediaries while transferring funds in order to evade detection of fraudulent activities. At the outset, you may not know how many intermediary groups there are or how long each path is from *x* to *y*. Each transaction generates a path of length *n*+1 in an *n*-intermediary transaction network. With a starting node *x* and a distance *n*, in order to establish the transaction flow, the weird flows pipeline verifies the maximum allowable transmission from node *x* to other nodes within that distance. The collection of all potential pathways that connect *x* and *y* up to a limit of *n* paths is defined as $Flow^n(x, y)$ according to equation (3).

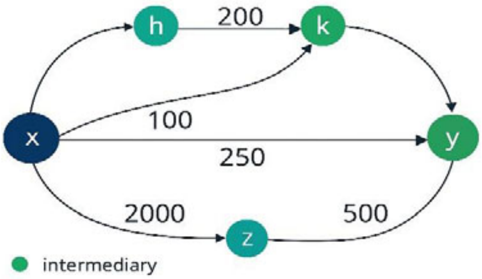
$$Flow(x, y) = \{P_1, ..., P_m\}, \text{ with } P_i \in Paths^n_{i,j}(G_T) \tag{5}$$

By minimising the weight of the links that link nodes *x* and *y*, the transaction network of node *x* specifies the maximum amount that node *y* may transmit to node *x* via intermediaries. For *x*-to-*y* transaction flows with multiple sets of intermediaries, the weight is the total of all the minimal weights. We state the maximum length *n* for the flow weight from *x* to *y* as:

$$W\left(Flow^*(x, y)\right) = \sum_{i=1}^{u} \min_{z_i \in paths^*(G_r)(Z_i)} (G_r)(Z_i), \text{ with } Z_i \in Paths^*(G_x) \tag{6}$$

Figure 2 shows a network that processes transactions. Beyond the edge weight ex, *y*, more factors must be considered in order to examine the hypothetical amount of money that is transferred from *x* to *y*. Consideration of intermediate nodes like *h*, *k*, and *z* is critical. The minimal minimum for each route is displayed in Table 1. The sum of all the minimal weights for all the pathways is this flow weight of 1,850 from *x* to *y*.

Figure 2 Transaction flow network showing intermediary nodes (h, k, z) between source node x and destination node y with weighted path values (see online version for colours)



Possible attempts to hide a more strongly weighted direct edge could be the *h*, *k*, and *z* paths that connect *x* and *y*. The user has Table 1. A two-column table is depicted in the image. 'P' and'min(W(P_i))' are the labels of the first and second columns, respectively. The table's rows detail several groupings of edges along with the lowest weights assigned to them.

Here is the content of the table:

- the smallest weight that may be applied to the set of edges $\{e(x, h), e(h, k), e(k, y)\}$ is 100
- all edges in the set $\{e(x, k), e(k, y)\}$ can be given a weight of zero
- a weight of 500 is the minimum for the set of edges $\{e(x, z), e(z, y)\}$
- an edge set $\{e(x, y)\}$ can have a minimum weight of 250.

All pathways from *x* to *y* with a maximum distance of 3 are listed in Table 1, along with their associated minimum weights. Keep in mind that a network's temporal aggregation during the interval *T* is just a rough estimate. Each edge along a valid route must satisfy the requirement. Every transaction is associated with a timestamp *t*, hence, *tei* < *tei* + 1.

Table 1 Minimum weights of different paths between nodes in the network

P_i	*Min* $(w(P_i))$
$\{e(x, h), e\ (h, k), e\ (k, y)\}$	100
$\{e(x, k), e\ (k, y)\}$	150
$\{e(x, z), e\ (z, y)\}$	200
$\{e(x, y)\}$	250

3.3 Optimised proposed hybrid model design

In contrast to earlier hybrid architectures, our model employs a parallel approach to merge spatial and temporal data using an MLP. For underrepresented groups, this layout makes it easier to detect and prevent fraud. To identify fraudulent charges on credit cards, it is necessary to build a hybrid optimised CLST model that incorporates SMOTE. By combining three distinct deep learning architectures, the model makes fraud detection more accurate. The use of a CNN analysis to extract geographical data is necessary for identifying trends in transaction patterns. The purpose of this approach is to find patterns in the data over time by analysing the interdependencies in sequences of monetary transactions using long short-term memory (LSTM). Minimal processing occurs at the MLP classification layer, which improves prediction accuracy by integrating spatial and sequential data. The model is able to handle class imbalance and get better performance results through the integration of SMOTE with hyperparameter change. CNN for Spatial Feature Extraction: CNNs are built using three main layers: convolutional, pooling, and fully connected.

These layers automatically adapt to new environments by means of backpropagation, allowing the network to learn spatial hierarchies (Mazumder et al., 2025). Neural networks, coupled with weights and biases that can be learned, make up the structure. Inside the structure, you'll find layers that are designed for CNNs and fully connected layers. Convolutional layers of a CNN take in data and use it to extract spatial features. Equation (2) describes the structure of a convolution process as follows:

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau. \tag{7}$$

For 2D input data (such as pictures or transaction details), the convolution procedure can be stated discretely as follows:

$$(I*)(i, j) = \sum\nolimits_m \sum\nolimits_n (i-m, j-n)\ K(m, n) \tag{8}$$

Where the structure of the input feature map is represented as follows:

The output feature map index is denoted by (*i*), while the kernel or filter is denoted by *K*. Forecasting time series with learned stochastic random forests (LSTMs): the efficacy of LSTM models has been demonstrated in several time series prediction applications, such as CCF detection. A wide variety of gates is at your disposal including input, output, and forget gates, among many others. Locating and taking into consideration temporal dependencies in sequential data are the primary goal of long short-term memory RNNs. A summary of the LSTM cell's equations is as follows:

$$f_t = \sigma\left(w_f . \left[h_{t-1}, x_t\right] + b_f\right) \tag{9}$$

$$i_t = \sigma\left(w_i . \left[h_{t-1}, x_t\right] + b_i\right) \tag{10}$$

$$C_t = \tanh\left(W_c \cdot \left[h_{t-1}, x_t\right] + b_c\right) \tag{11}$$

$$C_t = f_t . C_{t-1} + i_t \cdot C_t \tag{12}$$

$$0_t = \left(w_o . \left[h_{t-1}, x_t\right] + b_o\right) h_t \tag{13}$$

$$= o_t \cdot \tanh C_t \tag{14}$$

where time is represented by *xb*. At time *b* – 1, the concealed state is represented by *t hb*. The condition of the cell at time *t* is represented by **q***b*. The sigmoid activation function is represented by *σ*, and weights and biases are denoted by *W* and *b*, respectively. We integrated sequential and spatial data into a single layer to enhance data management. The MLP's dense layers are fed the combined results of the CNN and LSTM into a feature vector, which is used for the final prediction. As a result, the model can more accurately detect correlations that contribute to anti-fraud efforts. The output layer with full connection: the component of the dense layer responsible for output is responsible for classifying transactions as either legitimate or fraudulent. It uses many fully connected layers to handle the combined CNN and LSTM output. Network computing is done by the hidden layers, with prediction made by the input and output layers. It is possible to execute the following calculation for every neuron in a dense layer:

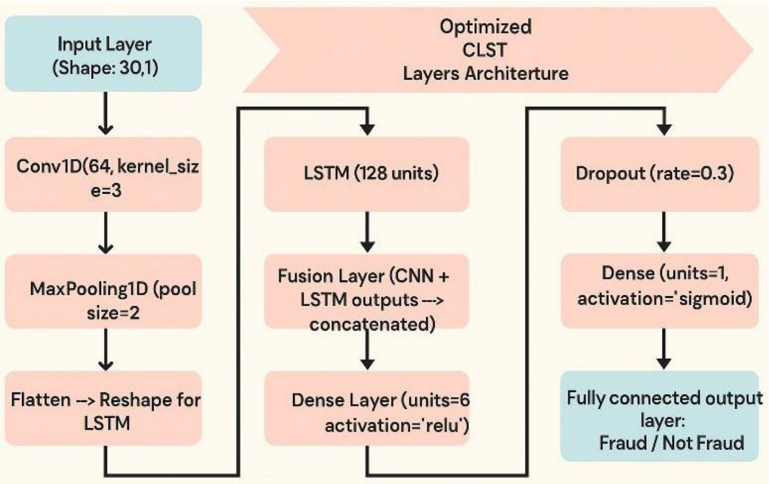
$$y = \left(\sum_{i=1}^{n} w_i x_i + b\right) \tag{15}$$

where input feature *xi* and weights *wi* are represented by the bias term is denoted by examples of activation functions are ReLU and sigmoid. In order to do binary classification, the last output layer makes use of a sigmoid activation function, which is supplied by:

$$\hat{y} = \sigma(W \cdot h + b) \tag{16}$$

where the projected probability of the positive class is represented by $\hat{b}$, and *h* is the output from the preceding hidden layer. In order to get the most out of each part, the best CLST architecture takes advantage of its unique capabilities. CNN layers successfully capture critical inter-feature interactions by extracting spatial characteristics and local patterns from transaction vectors. By modelling sequential dependencies across transactions, the LSTM layers are able to detect the temporal patterns of behaviour typically associated with fraudulent operations. Lastly, robust decision-making and high-level feature integration are made possible by a dense layer for categorisation that consists of numerous fully connected layers. Last but not least, this research employs a thick layer to convert the learnt feature representations into reliable fraud predictions. In Figure 3, we can see the CLST architecture that has been optimised.

Figure 3 Optimised CLST layer architecture (see online version for colours)



4 Results and discussion

4.1 Various models' comparative performance

The results of several models' combined performance are shown in Table 2. Regularly, our proposed feature extraction methods outperform state-of-the-art algorithms on measures such as Balanced Accuracy, G-Means, and F1-scores. These methods include SNN + RF and SNN + SVM. A higher false positive rate is a result of traditional anomaly detection algorithms' poor Precision and Specificity, despite their high Recall. This is in contrast to more modern methods like IF and OCSVM. While OCSVM and IF do a good job of catching abnormalities, our results show that their tradeoff in accuracy makes them impractical for uses where false alarms are expensive.

Table 2 Benchmarking model performance on five datasets

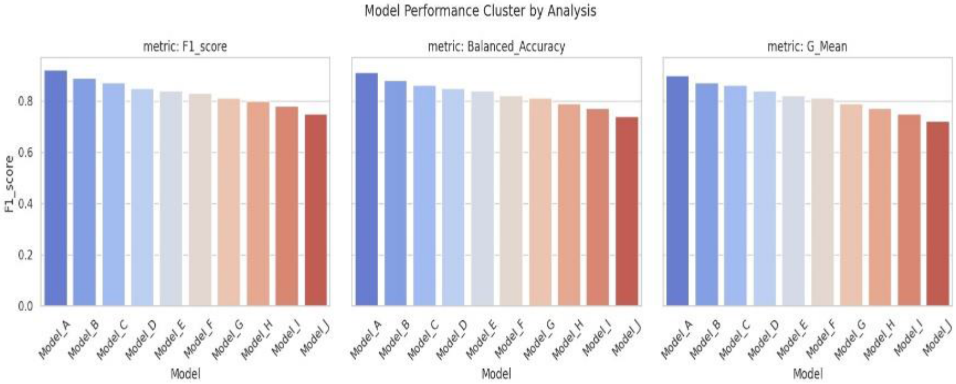
Model	*Precision*	*Recall*	*F1-score*	*G-mean*	*Balanced accuracy*	*Specificity*
IF	0.088 (0.125)	0.417 (0.480)	0.136 (0.191)	0.469 (0.412)	0.666 (0.209)	0.916 (0.092)
OCSVM	0.063 (0.069)	*0.841* (0.248)	0.111 (0.116)	0.646 (0.109)	0.674 (0.123)	0.508 (0.010)
RF	0.745 (0.430)	0.666 (0.397)	0.703 (0.412)	0.727 (0.414)	0.833 (0.198)	1.000 (0.000)
SVM	0.861 (0.173)	0.636 (0.349)	0.705 (0.340)	0.768 (0.239)	0.818 (0.174)	0.999 (0.001)
SMOTE + RF	0.736 (0.425)	0.719 (0.413)	0.727 (0.419)	0.757 (0.426)	0.859 (0.207)	0.999 (0.001)
SMOTE + SVM	0.727 (0.389)	0.771 (0.404)	0.583 (0.445)	0.819 (0.333)	0.880 (0.199)	0.989 (0.017)

Table 2 Benchmarking model performance on five datasets (continued)

Model	*Precision*	*Recall*	*F1-score*	*G-mean*	*Balanced accuracy*	*Specificity*
SNN	0.868 (0.229)	0.660 (0.417)	0.677 (0.398)	0.755 (0.332)	0.829 (0.208)	0.999 (0.001)
SNN + RF	*0.937* (0.087)	0.741 (0.372)	0.765 (0.338)	0.821 (0.287)	0.870 (0.186)	0.999 (0.001)
SNN + SVM	0.791 (0.212)	0.839 (0.275)	*0.787* (0.231)	*0.902* (0.174)	*0.918* (0.137)	0.997 (0.001)

We further explored the model's performance ranking and its statistical significance by employing Scott-Knott clustering analysis, utilising F1-score, G-mean, and balanced accuracy as the key metrics. As illustrated in Figure 4, the F1-score rankings demonstrate considerable variability. Traditional anomaly detection algorithms, SMOTE-based methods, and SNN-based models all rank well in this category (Vilella et al., 2025). G-Mean and Balanced Accuracy did not show any statistically significant differences amongst the models, suggesting that the models' overall ranking is consistent across these metrics, regardless of whether the algorithms in question have different capabilities.

Figure 4 Scott-Knott clustering outcomes for five datasets, evaluated with F1-score, balanced accuracy, and g-mean (see online version for colours)

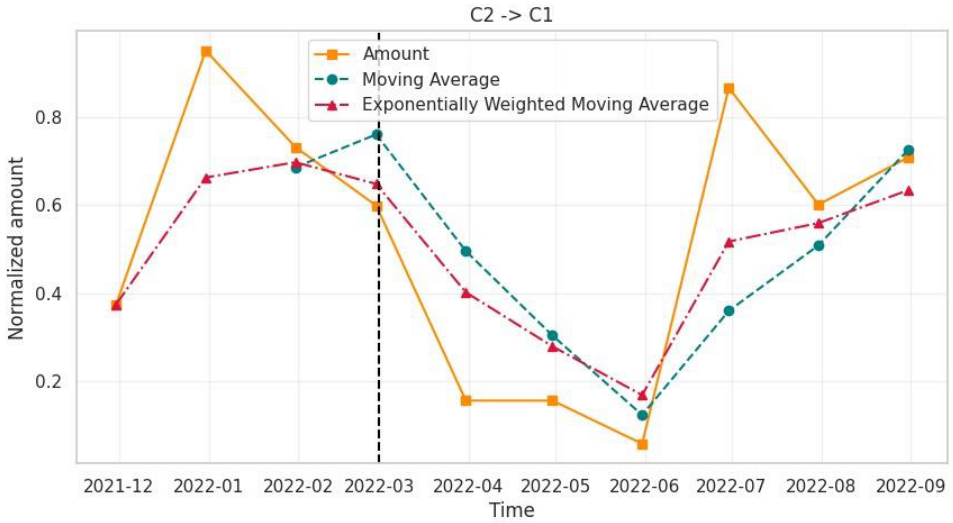


4.2 *Use case: as the conflict in Ukraine escalates, funds are moving between major European nations.*

The WeirdFlows tool was used in this example to investigate a financial crime. The example shows how WeirdFlows can help find complex patterns that could hide financial fraud using networks with different kinds of node aggregation. Following these procedures will ensure that your data is protected to the extent that the law and the AFC Digital Hub consortium have mandated. For example, after researchers receive their anonymised BIC codes, the amounts are adjusted to the highest value in the time series. To further ensure the data's security, the data provider has requested that all nation names be anonymised as C1, and C2. In this use case, we take a macro view of the financial flows between two major European nations, examining the time from the start of the conflict in C3 and the EU's economic sanctions. Figure 2 shows the weekly totals transferred directly from C2 BICs to C1 BICs. At the point where the grey dashed line appears in C3, the battle has begun. The economic fines did not affect the number of transactions from C2 to C1 because the pattern remained constant. To find the maximum data transfer rate from C2 to C1 BICs via a single intermediate, one can look at Figure 5, which illustrates the time series of w(Flow3(C2, C1)).

Peaks rather than patterns can be seen in both datasets, particularly before the war's commencement. An AFC analyst can look into the Flow3(C2, C1) intermediaries to find out more about the problem. The maximum length of the 86 pathways connecting C2 and C1 is 3. You can see which intermediaries' real value has increased the most compared to the predicted moving average after February 24 in Table 3. Among all the transactions that pass through C13, one jumps out with a 66% increase: country C13. Based on the methods outlined in [reference], the BIC identified as BIC03C2, associated with 4, was detected through the application of the anomaly identification pipeline on financial graphs. Take advantage of this foreign data. If AFCDigital Hub requests it, the data used to support the study's conclusions can be obtained from the ISP. Attention: Data availability limits are in effect. A non-disclosure agreement will be requested of researchers who wish to access the data for academic reasons.

Figure 5 Direct transfers between C2 and C1 BICs (see online version for colours)



Note: On 24 February 2022, the conflict in Ukraine began, as seen by the vertical grey dotted line.

It was accurately determined that country C2 was engaged in malevolent activities. After the war in Ukraine begins, the transaction flow becomes heavier, reaching a high in May and June (Figure 6). You can see this pattern when you use it as a starting point for WeirdFlows and examine all transactions passing via country C13 that include BIC03C2 and country C1.

Figure 6 Time series showing fluctuating amounts with smoothed 80-day and exponentially weighted moving averages revealing underlying trends amid high volatility (see online version for colours)

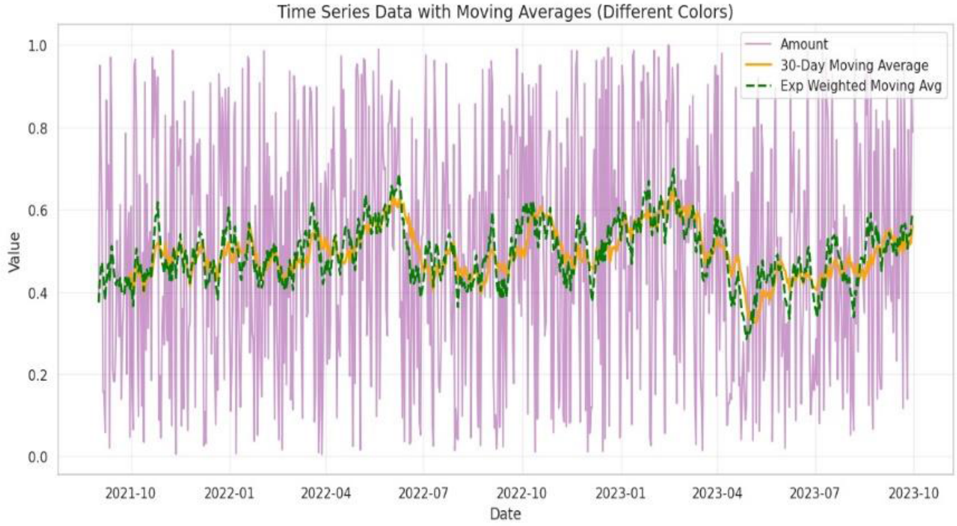


Figure 6 show the relevance of the revenue stream Flow3(C2, C1), where C2 and C1 are two nodes in a network, and C1 is the first node in the network. Table 3 Intermediaries in Flow3 (C2, C1) showing the highest percentage growth between the actual value and the expected value, as calculated by the moving average. Figure 7 shows the flow of transactions from C2 to C1 via C4, whereas Figure 8 shows the same flow via C5.

Table 3 Country-wise difference values

Country	*Difference*
C4	0.427
C6	0.436
C5	0.493
C7	0.535
C8	0.539
C9	0.549
C10	0.604
C11	0.636
C12	0.662
C13	0.665

Figures 7 and 8 display the flow weights calculated for aggregated networks every week on the left side. The right side illustrates the individual edges of the flow.

Figure 7 Funds being transferred from C2 to C1 through C4 (see online version for colours)

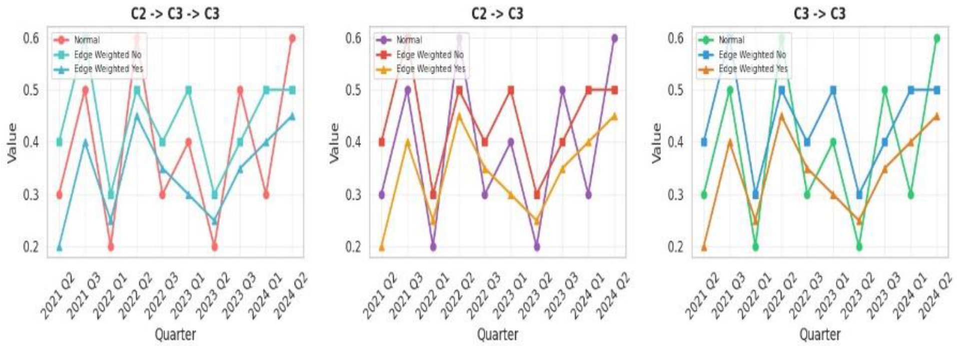


Figure 8 Funds being transferred from C2 to C1 through C5 (see online version for colours)

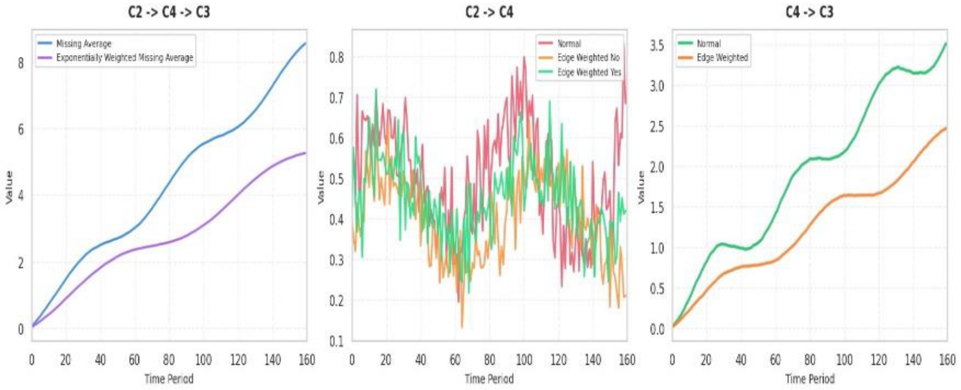
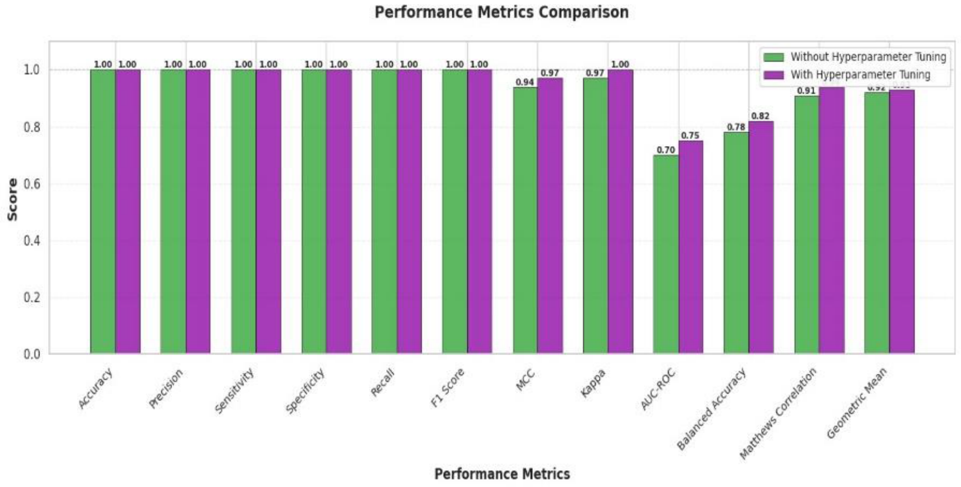


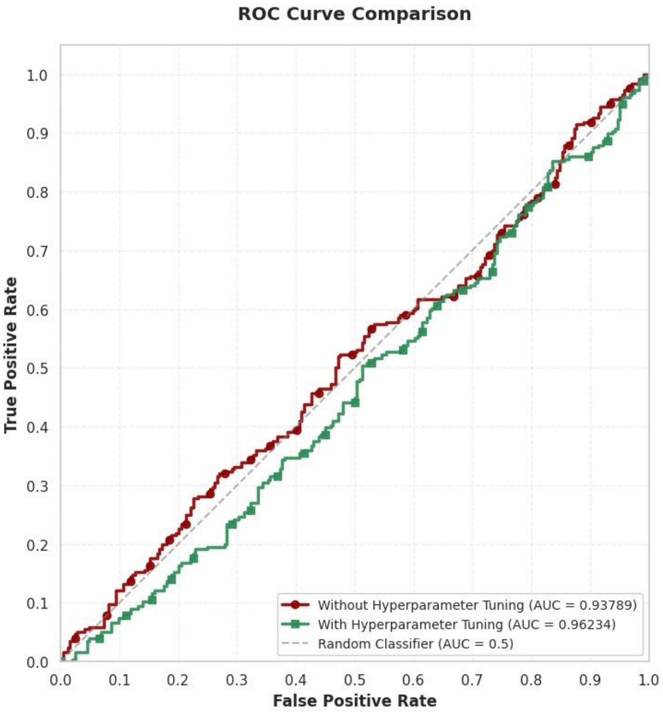
Figure 9 Improved model efficiency through hyperparameter optimisation (see online version for colours)



They are shown in Figures 4 and 5, respectively, as C2–C4 and C4–C1, and C5–C1 and C2–C1. The w(Flow3(C2, C1)) weight begins to exhibit speedy development in May 2022. This growth can be better understood by dissecting the specific edges that WeirdFlows uncovered. By the way, it is well-documented that very similar patterns of intermediaries have been employed to avoid international sanctions5. Important performance indicators were compared before and after hyperparameter adjustment in the bar chart. Improved recall, precision, and F1-score – measures for identifying fraud – were particularly brought about by the modified model (orange bars), as shown in Figure 9.

Figure 10 displays the ROC-AUC curve and explains its importance. One way to visualise the tradeoff between TPR and FPR is using the ROC curve (Jabeen, 2025). The model's performance is outstanding if the AUC is close to 1.0.

Figure 10 ROC-AUC curve for the CLST model evaluation (see online version for colours)



5 Conclusions

This work shows that in complicated real-time contexts, anomaly transaction detection and early-warning skills can be significantly improved by combining multisource financial data with advanced deep learning architectures. Optimal CLST offers better Recall, precision, and F1-scores than traditional and unimodal approaches by integrating spatial-temporal feature extraction through CNN, LSTM, and MLP layers and resolving data imbalance using SMOTE. This methodology may detect concealed transaction patterns and intermediate networks that might indicate fraudulent or sanction-evasion activities; the use case on cross-border financial flows during geopolitical events shows this in action. Strong, scalable, and context-aware anomaly detection are features offered by the framework, which the results demonstrate may be used to control systems in the financial and industrial sectors. In high-risk financial ecosystems, this strategy improves accuracy in fraud prevention, operational resilience, compliance, and strategic decision-making.

Declarations

All authors declare that they have no conflicts of interest.

References

Blazek, R. and Duricova, L. (2025) 'Beyond expectations: anomalies in financial statements and their application in modeling', *Stats*, Vol. 8, p.63, https://doi.org/10.3390/stats8030063.

Chidibere, J. (2024) *Multimodal Anomaly Detection: Combining Data from Different Sources*, ResearchGate Link.

Chou, E.P. (2025) 'Enhancing anomaly detection in structured data using Siamese neural networks as a feature extractor', *Mathematics*, Vol. 13, No. 7, p.1090.

Durana, P. (2022) 'The use of Beneish M-scores to reveal creative accounting: evidence from Slovakia', *Equilibrium: Quarterly Journal of Economics and Economic Policy*, Vol. 17, No. 12, pp.481–510.

Guo, H. (2021) 'Logbert: log anomaly detection via BERT', *Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN)*, Shenzhen, China, 18–22 July, pp.1–8.

Hanchuk, D.O. and Semerikov, S.O. (2025) 'Implementing MLOps practices for effective machine learning model deployment: a meta synthesis', *CEUR Workshop Proceedings*, pp.329–337.

He, S. (2021) 'A survey on automated log analysis for reliability engineering', *ACM Computing Surveys*, Vol. 54, No. 1, p.37, https://doi.org/10.1145/3460345.

Herden, O. (2020) 'Architectural patterns for integrating data lakes into data warehouse architectures', *Big Data Analytics: 8th International Conference, BDA 2020*, Sonepat, India, 15–18 December, Proceedings 8, pp.12–27, Springer International Publishing.

Jabeen, M. (2025) 'Enhanced credit card fraud detection using deep hybrid CLST model', *Mathematics*, Vol. 13, No. 12, p.1950.

Koay, A.M. (2023) 'Machine learning in industrial control system (ICS) security: current landscape, opportunities, and challenges', *Journal of Intelligent Information Systems*, Vol. 60, No. 2, pp.377–405.

Lee, M. (2023) *A Survey on the Real-world Cyberattacks on the Industrial Internet of Things*, Authorea Preprints.

Liu, Y. (2024) 'Temporal logical attention network for log-based anomaly detection in distributed systems', *Sensors*, Vol. 24, No. 24, p.7949, https://doi.org/10.3390/s24247949.

Manafi, H. (2025) 'An unsupervised fusion strategy for anomaly detection via Chebyshev graph convolution and a modified adversarial network', *Biomimetics*, Vol. 10, p.245, https://doi.org/10.3390/biomimetics10040245.

Mazumder, M.T.R. et al. (2025) 'Anomaly detection in financial transactions using convolutional neural networks', *Journal of Economics, Finance and Accounting Studies*, Vol. 7, No. 2, pp.195–207.

Mikuni, V. (2024) 'High-dimensional and permutation invariant anomaly detection', *SciPost Physics*, Vol. 16, No. 3, p.62.

Palakurti, N.R. (2024) 'Challenges and future directions in anomaly detection', in *Practical Applications of Data Processing, Algorithms, and Modeling*, pp.269–284, IGI Global, USA.

Qu, X. and Liu, Z. (2024) 'MFGAN: multimodal fusion for industrial anomaly detection using attention-based autoencoder and generative adversarial network', *Sensors*, Vol. 24, p.637, https://doi.org/10.3390/s24020637.

Rella, B.P.R. (2022) 'MLOPs and DataOps integration for scalable machine learning deployment', *International Journal for Multidisciplinary Research*, Vol. 4, Nos. 1–3, pp.2582–2160.

Riegler, M. (2021) 'Multi-mode systems for resilient security in Industry 4.0', *Procedia Computer Science*, Vol. 180, pp.301–307.

Sabău, A.I. (2021) 'A statistical model of fraud risk in financial statements: case for Romanian companies', *Risks*, Vol. 9, No. 6, p.116.

Shiva, K. (2024) 'Anomaly detection in sensor data with machine learning: predictive maintenance for industrial systems', *Journal of Electrical Systems*, Vol. 20, No. 10, pp.454–462.

Urdaneta-Camacho, R. and Guevara-Pérez, J.C. (2022) 'The other side of the "league of stars": analysis of the financial situation of Spanish football', *International Journal of Financial Studies*, Vol. 11, No. 1, p.3.

Vilella, S. et al. (2025) 'Weirdnodes: centrality-based anomaly detection on temporal networks for the anti-financial crime domain', *Applied Network Science*, Vol. 10, No. 1, pp.1–29.

Wu, H. and Xu, J. (2021) 'Autoformer: decomposition transformers with auto-correlation for long-term series forecasting', *Advances in Neural Information Processing Systems*, Vol. 34, No. 12, pp.22419–22430.

Xie, Y. (2020) 'An attention-based GRU network for anomaly detection from system logs', *IEICE Transactions on Information and Systems*, Vol. 103, pp.1916–1919, https://doi.org/10.1587/transinf.2020EDL8016.

Xu, J. and Wu, H. (2021) *Anomaly Transformer: Time Series Anomaly Detection with Association Discrepancy*, arXiv, arXiv:2110.02642.

Xu, L. and Shang, K. (2025) 'Multi-scale feature fusion-based real-time anomaly detection in industrial control systems', *Electronics*, Vol. 14, p.1645, https://doi.org/10.3390/electronics14081645.

Zhou, T. and Ma, Z. (2022) 'Fedformer: frequency-enhanced decomposed transformer for long-term series forecasting', *Proceedings of the International Conference on Machine Learning*, PMLR, Baltimore, MD, USA, 17–23 July, pp.27268–27286.