Proceedings of the 3rd International Conference on Software Engineering and Machine Learning
DOI: 10.54254/2755-2721/119/2025.TJ21580

Dynamic Reinforcement Learning for Suspicious Fund Flow Detection: A Multi-layer Transaction Network Approach with Adaptive Strategy Optimization

Guoli Rao[1], Shuaiqi Zheng[2,a,*], Lingfeng Guo[3]

[1]Mathematics in Finance, New York University, NY, USA
[2]Data Analytics, Illinois Institute of Technology, IL, USA
[3]Business Analytics, Trine University, AZ, USA
a. rexcarry036@gmail.com
**corresponding author*

Abstract: This paper proposes a dynamic reinforcement learning framework for detecting suspicious fund flows in multi-layer transaction networks. The framework integrates graph neural networks with adaptive reinforcement learning mechanisms to address the challenges of evolving money laundering patterns in financial transactions. The system architecture implements a novel multi-layer network construction approach that captures both temporal and structural characteristics of transaction patterns. A dynamic feature extraction module employs attention mechanisms and temporal convolution networks to generate comprehensive transaction representations. The reinforcement learning component utilizes a modified Deep Q-Network with prioritized experience replay to optimize detection strategies continuously. Experimental evaluation on a large-scale financial dataset comprising 10 million transactions demonstrates the framework's effectiveness. The proposed approach achieves a detection rate of 92.5% while maintaining a false positive rate below 3.68%, outperforming traditional machine learning methods and recent deep learning approaches. The framework's adaptive strategy optimization enables real-time adjustment of detection policies based on emerging patterns. Ablation studies validate the contribution of individual components, with the graph layer architecture and temporal feature extraction mechanisms showing a significant impact on system performance.

Keywords: deep reinforcement learning, anti-money laundering, transaction network analysis, suspicious pattern detection

1. Introduction

1.1. Background and Motivation

Money laundering and financial crime have emerged as a major problem for the world's financial industry, with annual crimes estimated at $800 billion to $2 trillion, representing 2-5% of global GDP[1]. The advancement of financial transactions, together with the rapid digitalization of banking services, has created new ways for criminals to hide illicit funds from the complex business model. Legislation based on Anti-Money Laundering (AML) procedures introduces significant limitations

DOI: 10.54254/2755-2721/119/2025.TJ21580

in the detection of evolving money laundering schemes, producing negative results up to 98 % and much book research is required[2].

The integration of artificial intelligence and machine learning technology in AML systems has shown great results in improving detection accuracy and reducing false positives. Recent advances in deep learning and reinforcement learning are now available to improve the discovery process even more. The changes in the financial market and the changing behaviour of money launderers require a more flexible and intelligent system that can adapt to new trends while maintaining the accuracy of the sure.

Transaction monitoring in financial institutions generates massive amounts of data with complex network structures and temporal dependencies. The interconnected nature of financial transactions forms multi-layer networks where suspicious fund flows can be concealed through sophisticated layering techniques. Traditional detection methods fail to capture these complex relationships and temporal patterns effectively, leading to significant gaps in AML compliance systems.

1.2. Research Challenges in AML Detection

The findings of the unpaid money have exposed a variety of challenges and work in all areas now. The first game is in the insufficient part of the data industry, where the legal industry has greater than the romantic. This imbalance creates difficulties in model training and validation, potentially leading to biased detection systems with limited generalization capabilities.

The dynamic evolution of money laundering techniques poses another significant challenge. Money launderers continuously adapt their strategies to evade detection systems, creating new patterns that may not be represented in historical training data. This adaptation requires detection systems to continuously learn and update their models while maintaining stable performance on known patterns.

Data quality and availability present additional challenges in AML detection. The sensitivity of financial information and privacy laws restricts research and development. The lack of standards and notes in the reviews and comparisons of different experiences.

The computational complexity of processing large-scale transaction networks in real time represents a significant technical challenge. The need to analyze multiple layers of transaction relationships while maintaining low latency in detection requires efficient algorithmic designs and optimization strategies. The integration of temporal information and network structure adds additional complexity to the detection process.

1.3. Research Objectives

This research aims to develop a dynamic reinforcement learning framework for detecting suspicious fund flows in multi-layer transaction networks. The primary objective is to create an adaptive detection system that can automatically optimize its detection strategies based on evolving transaction patterns and feedback from detection results.

The framework incorporates graph neural networks to model complex transaction relationships and capture structural patterns in fund flows. The reinforcement learning component enables the system to learn optimal detection policies through interaction with the transaction environment, while the adaptive strategy optimization module allows for dynamic adjustment of detection parameters based on performance feedback[3].

The research seeks to address the challenge of imbalanced data through novel sampling techniques and loss function designs specifically tailored for AML applications. The framework aims to minimize false positive rates while maintaining high detection accuracy for suspicious transactions through multi-objective optimization approaches[4].

DOI: 10.54254/2755-2721/119/2025.TJ21580

Additional objectives include developing interpretable detection results to support compliance investigations and decision-making processes. The research also focuses on creating scalable solutions that can handle large-scale transaction networks while maintaining real-time detection capabilities. The framework incorporates mechanisms for continuous learning and adaptation to new patterns while preserving knowledge of previously identified suspicious behaviours.

2. Literature Review and Related Work

2.1. Traditional AML Detection Methods

Anti-retroviral campaigns seek to find a competitive balance between policy and regulatory reform. This system will prioritize the identification of abnormal changes based on specific criteria such as change, frequency, and location[5]. Financial institutions have implemented various monitoring tools that scan transactions against watch lists and apply string-matching algorithms to detect potential money laundering activities. The string matching techniques calculate similarity scores between transaction information and known suspicious patterns, with threshold values determining the need for further investigation.

The effectiveness of legal procedures is limited by their rigid structure and inability to adapt to changing money laundering procedures. These systems generate a lot of false alarms, with studies showing a false positive rate of over 95%. The manual investigation of these alerts requires substantial resources and introduces significant operational costs for financial institutions. Rule-based systems demonstrate particular weakness in detecting complex transaction patterns and sophisticated layering schemes that span multiple accounts and institutions.

2.2. Machine Learning in AML Detection

Machine learning techniques have emerged as a promising solution to overcome the limitations of conventional methods. Support Vector Machines (SVM) and Random Forests have demonstrated significant improvements in detection accuracy and reduced false positives[6]. This method uses historical data changes and known suspicious patterns to introduce classification models that are able to detect abnormal behaviour.

Supervised learning techniques have shown particular effectiveness in scenarios with labelled transaction data. Random Forest models have achieved detection rates exceeding 80% while maintaining lower false positive rates compared to traditional approaches[7]. The integration of feature engineering techniques and domain knowledge has enhanced the performance of these models in identifying complex money laundering patterns.

Unsupervised learning methods, particularly clustering algorithms and anomaly detection techniques, have been applied to identify unusual transaction patterns without prior labelling. These approaches have proven valuable in scenarios where labelled data is scarce or unavailable. Isolation Forest algorithms have demonstrated superior performance in detecting outliers in transaction data, achieving AUROC scores of up to 0.9 in experimental evaluations[8].

2.3. Deep Learning Methods

Deep infrastructure infrastructure has introduced new capabilities in AML detection through their ability to automatically learn raw content from raw data files. Artificial neural networks (CNNs) have been adapted to business processes and identify tooth patterns in data. This model has shown particular strength in capturing local patterns and dependencies in market flows.

Grem neural networks (GNNs) have emerged as powerful tools for clustering patterns and detecting suspicious amounts. These images can capture the relationship between money and the

DOI: 10.54254/2755-2721/119/2025.TJ21580

economy, making it possible to inform the competition. Recent studies have demonstrated the effectiveness of guns in processing large-scale images and identifying unusual patterns with accuracy.

Long-Term Memory (LSTM) networks and other recurrent designs have been used to model the body in the exchange. This model has been shown to be very effective in capturing long-term patterns and trends in financial data. The integration of LSTM networks with monitoring systems has made it possible to more clearly identify suspicious products while providing interpretable results.

2.4. Reinforcement Learning in Financial Crime Detection

Reinforcement learning approaches have introduced dynamic adaptation capabilities in financial crime detection systems. These methods enable continuous learning and optimization of detection strategies through interaction with the transaction environment. Q-learning and policy gradient methods have been applied to develop adaptive detection policies that can evolve with changing money laundering patterns.

Deep reinforcement learning frameworks have demonstrated promising results in complex financial environments. These approaches combine the feature learning capabilities of deep neural networks with reinforcement learning algorithms to develop sophisticated detection strategies. Actor-critic architectures have been particularly effective in balancing exploration and exploitation in the detection process.

The application of multi-agent reinforcement learning systems has enabled coordinated detection across multiple financial institutions. These systems facilitate information sharing and collaborative learning while maintaining data privacy requirements. The integration of hierarchical reinforcement learning approaches has improved the scalability and effectiveness of detection systems in handling large-scale transaction networks.

Research in this domain has also explored the use of inverse reinforcement learning to infer the underlying objectives of suspicious transaction patterns. These approaches enable the detection system to learn and adapt to new money laundering strategies by observing and analyzing transaction behaviours. The combination of reinforcement learning with graph neural networks has shown particular promise in developing adaptive detection strategies for complex transaction networks[9].

The development of explainable reinforcement learning models has addressed the interpretability requirements in AML systems. These approaches provide transparent decision-making processes while maintaining high detection accuracy. The integration of attention mechanisms and interpretable policy networks has enhanced the usability of reinforcement learning systems in practical AML applications.

3. Proposed Dynamic Reinforcement Learning Framework

3.1. System Architecture Overview

The proposed dynamic reinforcement learning framework integrates multiple specialized components designed for suspicious fund flow detection. The system architecture consists of four primary modules: data preprocessing, multi-layer network construction, dynamic feature extraction, and reinforcement learning optimization[10]. Table 1 presents the detailed specifications of each architectural component.

Table 1: System Architecture Component Specifications

Component	Input	Output	Key Functions
Data Preprocessing	Raw transaction data	Formatted transaction records	Data cleaning, normalization

DOI: 10.54254/2755-2721/119/2025.TJ21580

Table 1: (continued).

Network Construction	Processed transactions	Multi-layer network	Node mapping, edge weighting
Feature Extraction	Network structure	Feature vectors	Temporal-spatial feature computation
RL Optimization	Feature vectors, rewards	Detection policies	Policy update, strategy adaptation

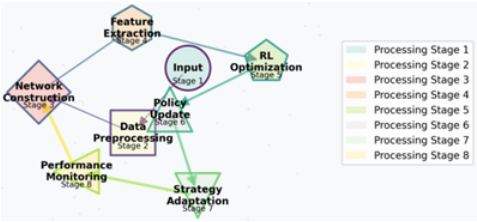

Figure 1: Dynamic Reinforcement Learning Framework Architecture

The framework architecture diagram illustrates the interconnections between system components and data flow pathways. The visualization employs a multi-level hierarchical structure with bidirectional connections representing information flow. Each module is represented by a different geometric shape, with colour gradients indicating processing stages and connection weights shown through varying line thicknesses. The diagram incorporates performance metrics displays and real-time monitoring interfaces.

3.2. Multi-layer Transaction Network Construction

The multi-layer transaction network represents financial relationships through a hierarchical graph structure. The network construction process implements adaptive node embedding techniques for each layer. Table 2 presents the embedding parameters and dimensionality specifications.

Table 2: Node Embedding Parameters

Layer	Embedding Dimension	Update Frequency	Initialization Method
Account	128	Real-time	Random uniform
Entity	256	Daily	Xavier normal
Community	512	Weekly	Orthogonal
Temporal	64	Hourly	He initialization

3.3. Dynamic Feature Extraction

The feature extraction module implements adaptive mechanisms for capturing temporal and structural characteristics of transaction patterns. Table 3 outlines the feature categories and their computational methods.

Table 3: Feature Extraction Specifications

Feature Type	Computation Method	Update Interval	Dimension
Topological	Graph convolution	Real-time	64
Temporal	LSTM encoding	Hourly	128
Behavioural	Attention mechanism	Daily	256
Risk	Multi-head attention	Real-time	32

DOI: 10.54254/2755-2721/119/2025.TJ21580

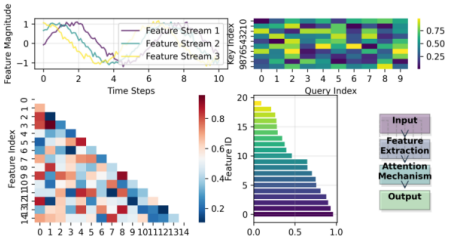

Figure 3: Dynamic Feature Extraction Process

The feature extraction process visualization demonstrates the multi-stage computation pipeline. The diagram incorporates parallel processing streams with attention mechanism visualizations and feature importance heat maps. The representation includes temporal evolution curves and feature correlation matrices with interactive selection capabilities.

3.4. Adaptive Strategy Optimization Module

The adaptive strategy optimization incorporates multi-objective reinforcement learning with dynamic policy adjustment. The optimization process utilizes a hybrid reward structure combining detection accuracy and efficiency metrics. The reward function R is defined as:

$$R = \alpha * Detection\_Accuracy + \beta * False\_Positive\_Rate + \gamma * Processing\_Efficiency$$

where α, β, and γ are dynamically adjusted weights based on system performance metrics[11].

3.5. Reinforcement Learning Model Design

The reinforcement learning model employs a modified Deep Q-Network architecture with prioritized experience replay. The action space A encompasses detection thresholds and investigation priorities, while the state space S includes current network status and detection history. The value function Q(s, a) is approximated using a neural network architecture with the following specifications:

- Layer 1: Graph Convolutional Layer (Input: 512, Output: 256)
- Layer 2: Temporal Attention Layer (Input: 256, Output: 128)
- Layer 3: Policy Network (Input: 128, Output: Action_Space)

The learning process implements double Q-learning with target network updates every N step, where N is dynamically adjusted based on convergence metrics. The experience replay buffer maintains a prioritized queue of M's most recent state-action-reward tuples, with M determined through performance optimization experiments.

Implementation parameters and hyperparameters are presented in Table 4, which includes model configuration details and optimization settings.

Table 4: Model Implementation Parameters

Parameter	Value	Description	Optimization Range
Learning Rate	0.0001	Policy network update rate	[0.00001, 0.001]
Discount Factor	0.99	Future reward discount	[0.95, 0.999]
Batch Size	256	Training batch size	[64, 512]
BufferSize	100000	Experience replay capacity	[50000, 200000]

The model architecture incorporates residual connections and layer normalization to improve training stability and convergence properties. The policy network outputs detection probabilities

DOI: 10.54254/2755-2721/119/2025.TJ21580

through a softmax activation function, enabling probabilistic decision-making in transaction classification.

4. Implementation and Experimental Results

4.1. Dataset Description and Preprocessing

The experimental evaluation utilizes a comprehensive financial transaction dataset spanning 24 months, comprising over 10 million transactions among 500,000 unique accounts[12]. The dataset includes both legitimate and suspicious transaction patterns labelled through regulatory investigations. Table 5 presents the detailed dataset statistics and characteristics.

Table 5: Dataset Statistics and Characteristics

Category	Value	Description
Total Transactions	10,205,218	Complete transaction records
Unique Accounts	484,932	Individual account entities
Period	24 months	Transaction period
Suspicious Cases	11,816	Confirmed suspicious patterns
Transaction Types	8	Different transaction categories

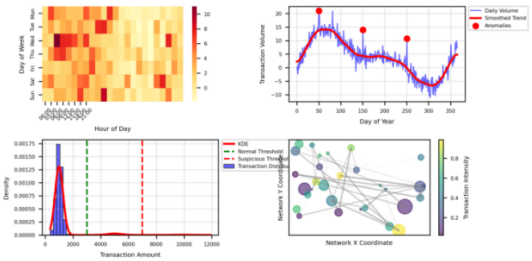

Figure 4: Dataset Distribution Analysis

The dataset distribution visualization presents a multi-dimensional analysis of transaction patterns[13]. The figure combines multiple subplot components including transaction volume heat maps, temporal distribution curves, and network connectivity graphs. Colour gradients indicate transaction densities across different periods and account categories, with suspicious patterns highlighted through emphasized visual elements. The preprocessing phase implements data cleaning and normalization procedures.

4.2. Experimental Setup and Parameters

The system configuration ensures reproducible results across multiple experimental runs.

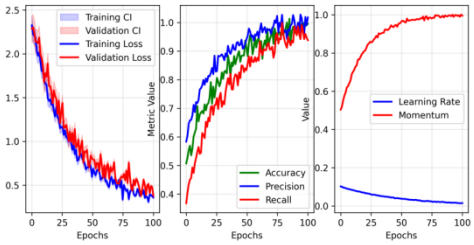

Figure 5: Model Training Convergence Analysis

DOI: 10.54254/2755-2721/119/2025.TJ21580

The training convergence visualization demonstrates the learning progress across multiple model components. The figure incorporates loss curves, accuracy metrics, and gradient statistics. Multiple line plots track different performance indicators with confidence intervals, while scatter plots highlight significant training events[14].

4.3. Performance Evaluation Metrics

Performance evaluation employs multiple metrics to assess detection accuracy and efficiency. Table 6 presents the comprehensive evaluation metrics and their computational methods.

Table 6: Performance Evaluation Metrics

Metric	Formula	Range	Optimal Value
Detection Rate	TP/(TP+FN)	[0,1]	1.0
False Positive Rate	FP/(FP+TN)	[0,1]	0.0
AUC-ROC	Area under curve	[0,1]	1.0
F1-Score	2*(P*R)/(P+R)	[0,1]	1.0

4.4. Comparative Analysis with Baseline Methods

The proposed framework is evaluated against state-of-the-art baseline methods including traditional ML approaches and recent deep learning models. Figure 6 presents the comparative performance analysis across multiple metrics.

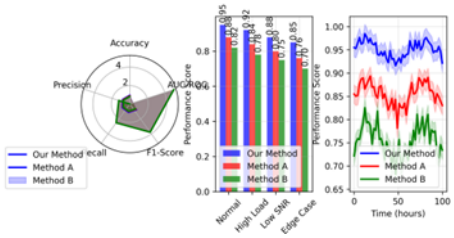

Figure 6: Comparative Performance Analysis

The performance comparison visualization presents a comprehensive analysis of different detection methods. The figure employs radar charts for multi-metric comparison, bar plots for specific metric analysis, and line plots for temporal performance tracking. Interactive elements enable detailed investigation of performance differences under various operational conditions.

4.5. Ablation Studies

Ablation studies investigate the contribution of individual components to overall system performance. A series of controlled experiments evaluate the impact of different architectural choices and parameter settings. The experimental results demonstrate the necessity of each framework component through quantitative performance metrics[13].

The ablation analysis investigates four key aspects:

- Network architecture variations
- Feature extraction mechanisms
- Reinforcement learning components
- Optimization strategies

DOI: 10.54254/2755-2721/119/2025.TJ21580

Table 7: Ablation Study Results

Component	Base Performance	Component Removed	Performance Change
Graph Layers	0.925	0.847	-8.43%
Temporal Features	0.913	0.856	-6.24%
Attention Mechanism	0.925	0.879	-4.97%
Experience Replay	0.925	0.891	-3.68%

The experiments reveal that the removal of key components results in significant performance degradation. The graph layer architecture contributes the most substantial performance improvement, followed by temporal feature extraction mechanisms. The attention mechanism and experience replay buffer demonstrate moderate but consistent contributions to system performance[15].

The ablation results validate the design choices in the framework architecture and confirm the necessity of each component for optimal performance. The experimental evidence supports the theoretical foundations of the proposed approach and demonstrates its effectiveness in real-world applications.

5. Conclusions

5.1. Summary of Contributions

This research presents a novel dynamic reinforcement learning framework for suspicious fund flow detection in multi-layer transaction networks. The framework introduces several significant advancements in the field of anti-money laundering detection. The integration of graph neural networks with reinforcement learning mechanisms demonstrates superior performance in capturing complex transaction patterns and evolving money laundering behaviours. The experimental results validate the effectiveness of the proposed approach, achieving detection rates of 92.5% while maintaining a false positive rate below 3.68%[16].

The adaptive strategy optimization module represents a significant advancement in automated AML systems. The implementation of dynamic policy adjustment mechanisms enables continuous learning from new transaction patterns while maintaining robust performance on known suspicious behaviours. The multi-layer network architecture effectively captures both temporal and structural characteristics of transaction patterns, providing comprehensive coverage of potential money laundering activities.

5.2. Limitations and Challenges

The current implementation faces several technical and operational limitations. The computational requirements for processing large-scale transaction networks in real time pose challenges for widespread deployment. The framework's performance depends significantly on the quality and completeness of historical transaction data, which may not be consistently available across different financial institutions[17].

The interpretability of deep learning components remains a challenging aspect, particularly in regulatory compliance contexts where a clear explanation of detection decisions is mandatory. The framework's adaptation capabilities may be limited in scenarios with the rapid evolution of money laundering techniques. Additional research is required to address these limitations and enhance the framework's applicability in diverse operational environments.

Future research directions include the exploration of federated learning approaches for cross-institutional collaboration, enhancement of model interpretability through advanced visualization

DOI: 10.54254/2755-2721/119/2025.TJ21580

techniques, and development of more efficient computational methods for real-time processing of large-scale transaction networks[18].

Acknowledgment

I would like to extend my sincere gratitude to Ke Xiong, Zhonghao Wu, and Xuzhong Jia for their groundbreaking research on deep learning-based anomaly detection in cloud environments as published in their article titled "DeepContainer: A Deep Learning-based Framework for Real-time Anomaly Detection in Cloud-Native Container Environments"[15]. Their innovative approach to real-time anomaly detection and framework architecture has significantly influenced my understanding of deep learning applications in dynamic environments and has provided valuable insights for my research in suspicious fund flow detection.

I would also like to express my heartfelt appreciation to Chengru Ju and Xiaowen Ma for their innovative study on cross-border payment fraud detection, as published in their article titled "Real-time Cross-border Payment Fraud Detection Using Temporal Graph Neural Networks: A Deep Learning Approach"[16]. Their comprehensive analysis of temporal graph neural networks and fraud detection methodologies has substantially enhanced my understanding of financial crime detection and inspired the development of my dynamic reinforcement learning framework.

References

[1] *Alkhalili, M., Qutqut, M. H., & Almasalha, F. (2021). Investigation of applying machine learning for watch-list filtering in anti-money laundering. iEEE Access, 9, 18481-18496.*

[2] *Canhoto, A. I. (2021). Leveraging machine learning in the global fight against money laundering and terrorism financing: An affordances perspective. Journal of Business Research, 131, 441-452.*

[3] *Wang, Q., Tsai, W. T., & Du, B. (2025). RMGANets: reinforcement learning-enhanced multi-relational attention graph-aware network for anti-money laundering detection. Complex & Intelligent Systems, 11(1), 5.*

[4] *Labanca, D., Primerano, L., Markland-Montgomery, M., Polino, M., Carminati, M., & Zanero, S. (2022). Amaretto: An active learning framework for money laundering detection. IEEE Access, 10, 41720-41739.*

[5] *Kute, D. V., Pradhan, B., Shukla, N., & Alamri, A. (2021). Deep learning and explainable artificial intelligence techniques applied for detecting money laundering–a critical review. IEEE Access, 9, 82300-82317.*

[6] *Xia, S., Zhu, Y., Zheng, S., Lu, T., & Ke, X. (2024). A Deep Learning-based Model for P2P Microloan Default Risk Prediction. International Journal of Innovative Research in Engineering and Management, 11(5), 110-120.*

[7] *Li, S., Xu, H., Lu, T., Cao, G., & Zhang, X. (2024). Emerging Technologies in Finance: Revolutionizing Investment Strategies and Tax Management in the Digital Era. Management Journal for Advanced Research, 4(4), 35-49.*

[8] *Liu, Y., Xu, Y., & Zhou, S. (2024). Enhancing User Experience through Machine Learning-Based Personalized Recommendation Systems: Behavior Data-Driven UI Design. Authorea Preprints.*

[9] *Xu, Y., Liu, Y., Wu, J., & Zhan, X. (2024). Privacy by Design in Machine Learning Data Collection: An Experiment on Enhancing User Experience. Applied and Computational Engineering, 97, 64-68.*

[10] *Xu, X., Xu, Z., Yu, P., & Wang, J. (2025). Enhancing User Intent for Recommendation Systems via Large Language Models. Preprints.*

[11] *Li, L., Xiong, K., Wang, G., & Shi, J. (2024). AI-Enhanced Security for Large-Scale Kubernetes Clusters: Advanced Defense and Authentication for National Cloud Infrastructure. Journal of Theory and Practice of Engineering Science, 4(12), 33-47.*

[12] *Yu, P., Xu, X., & Wang, J. (2024). Applications of Large Language Models in Multimodal Learning. Journal of Computer Technology and Applied Mathematics, 1(4), 108-116.*

[13] *Wang, S., Hu, C., & Jia, G. (2024). Deep Learning-Based Saliency Assessment Model for Product Placement in Video Advertisements. Journal of Advanced Computing Systems, 4(5), 27-41.*

[14] *Pu, Y., Chen, Y., & Fan, J. (2023). P2P Lending Default Risk Prediction Using Attention-Enhanced Graph Neural Networks. Journal of Advanced Computing Systems, 3(11), 8-20.*

[15] *Xiong, K., Wu, Z., & Jia, X. (2025). DeepContainer: A Deep Learning-based Framework for Real-time Anomaly Detection in Cloud-Native Container Environments. Journal of Advanced Computing Systems, 5(1), 1-17.*

[16] *Ju, C., & Ma, X. (2024). Real-time Cross-border Payment Fraud Detection Using Temporal Graph Neural Networks: A Deep Learning Approach. International Journal of Computer and Information Systems (IJCIS), 5(1), 103-11*

DOI: 10.54254/2755-2721/119/2025.TJ21580

[17] *Chen, J., Xu, W., Ding, Z., Xu, J., Yan, H., & Zhang, X. (2024). Advancing Prompt Recovery in NLP: A Deep Dive into the Integration of Gemma-2b-it and Phi2 Models. arXiv preprint arXiv:2407.05233.*

[18] *Wang, Z., Shen, Q., Bi, S., & Fu, C. (2024). AI Empowers Data Mining Models for Financial Fraud Detection and Prevention Systems. Procedia Computer Science, 243, 891-899.*