Revija za Kriminalistiko in Kriminologijo
REVIJA ZA KRIMINALISTIKO IN KRIMINOLOGIJO,
ISSN :0034-690X,
STEFANOVA ULICA 2, LJUBLJANA,
SLOVENIA, 1501
Vol. 77 No. 1 (2026)
OPEN ACCESS

Research on Risk Identification and Early Warning Models for Fintech Crimes Based on Big Data Analysis

Yingjian Li[1], Liping Jia[2] *

[1] School of Economics, Shanxi University of Finance and Economics, Taiyuan, 030000, Shanxi, China

[2] School of Finance, Shanxi University of Finance and Economics, Taiyuan, 030000, Shanxi, China

***Email:**JiaLiping@tbu.edu.gr

Abstract

In response to the diverse evolution of fintech crimes—manifesting in behavioral pathways, identity spoofing, and cross-platform attacks—this study investigates a multi-stage risk identification and early warning model integrating big data analytics. The model achieves hierarchical identification of anomalous accounts through structural feature screening, deep modeling of temporal behaviors, and a Stacking ensemble strategy. It further incorporates a dynamic threshold mechanism and an online self-learning module to enhance adaptability. A testing platform under real business conditions was constructed. Comparisons with models such as GBDT, BiLSTM, and Transformer showed that the proposed model achieved an AUC of 0.941 and an F1-score of 0.893, with a significantly lower standard deviation than other methods, demonstrating strong stability and robustness.

Keywords

Fintech crime; Big data analysis; Risk identification; Ensemble model; Dynamic early warning

1 INTRODUCTION

The rapid advancement of fintech has driven profound transformations in financial service models while simultaneously fostering more covert, high-frequency, and sophisticated criminal activities. Existing risk control systems often rely on static rules or single models, struggling to adapt to risk environments characterized by interwoven multi-source data and dynamic attack paths. Building a multi-stage identification and early warning mechanism that integrates structural features and temporal behavior has become a key technological pathway for enhancing financial security prevention capabilities. This approach holds significant theoretical value and engineering significance for ensuring the resilience and compliant operation of the financial ecosystem.

2 ANALYSIS OF TYPES AND CHARACTERISTICS OF FINTECH CRIME

The stealth and technical nature of fintech crimes primarily stem from attackers' precise identification and rapid exploitation of structural vulnerabilities within financial systems. Criminals leverage asymmetric information advantages in user identity verification, transaction routing, and account behavior mapping on data-driven platforms to evade real-time detection of anomalous transactions by traditional regulatory tools. As financial services increasingly migrate to mobile platforms and distributed architectures, microservice frameworks interconnected via APIs have become technical springboards for cross-chain money laundering and identity hijacking. Building upon this, technology-driven financial crimes increasingly leverage high-frequency trading, disguised smart contracts, and cross-border anonymous payment tools. By fragmenting fund flows and signal paths, they construct complex transfer chains that compress the reaction window of early warning systems. Simultaneously, criminals embed adversarial perturbation signals into model learning

mechanisms, deliberately misleading intelligent risk control systems into misclassifying risk levels to circumvent and manipulate system boundary policies [1]. Compared to traditional financial crimes, fintech crimes exhibit accelerated attack rhythms, minimal data traces, and strong adaptive intelligence—evolutionary traits demanding higher timeliness and robustness from existing rule-based matching and static feature recognition models.

3 DESIGN OF MULTI-STAGE INTEGRATED RISK IDENTIFICATION MODEL

3.1 Principles and Framework of Multi-stage Ensemble Modeling

The multi-stage ensemble modeling framework leverages the distribution characteristics of risk signals across structural and sequential dimensions to construct a hierarchical modeling system centered on the "coarse identification—fine identification—ensemble decision" workflow (see Figure 1). The first stage employs gradient-boosted tree algorithms to establish structural feature screening and preliminary risk assessment mechanisms, identifying account behavior patterns significantly deviating from normal patterns. This stage generates credible initial risk values and sample distribution weights for subsequent stages [2]. The second stage employs a deep architecture integrating convolutional neural networks (CNNs) and gated recurrent neural networks (GRNNs) to jointly model continuous transaction sequences, operational trajectories, and time-correlated behavioral features, learning high-order patterns across time windows within the feature space. The training objective function for this stage is defined as follows:

$$L_{seq} = \frac{1}{T}\sum_{t=1}^{T}\left[y_t \log\left(\hat{y}_t\right) + \left(1-y_t\right)\log\left(1-\hat{y}_t\right)\right] + \lambda\left\|\theta\right\|_2^2 \quad (1)$$

where y_t is the true risk label at time step t , $\hat{y}_t$ is the model prediction probability, θ is the set of parameters to be optimized, and λ is the regularization coefficient to prevent model overfitting. The third stage introduces a meta-learning framework based on Stacking, integrating the outputs from the first two stages. It employs Logistic regression or shallow neural networks as fusion engines to dynamically weight submodel results and output the final risk score.

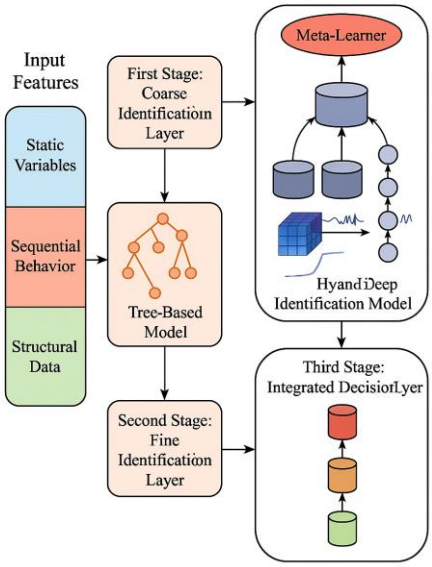

Figure 1 Overall Structure of the Multi-Stage Model

3.2 Stage 1: Coarse Identification Layer Based on Tree Models

The coarse identification layer in Phase 1 primarily relies on Gradient Boosting Decision Trees (GBDT) to construct a feature-driven preliminary risk screening model. This captures static variable patterns with significant discriminative power within financial transaction samples. This stage's model utilizes account attributes, registration behavior, device information, and basic transaction statistics as primary input features. It employs a cascade of weak classifiers to achieve iterative error correction, enhancing nonlinear boundary fitting capabilities while maintaining high computational efficiency [3]. During training, the model minimizes the weighted residual squared error loss function, whose objective can be expressed as:

$$L_{CBDT} = \sum_{i=1}^{n} \left[w_i y_i - F_m(x_i) \right]^2 \quad (2)$$

where `$n$` denotes the total number of samples, `$y_i$` represents the true label of the `$i$` th sample, `$F_m(x_i)$` is the cumulative prediction value of the `$m$` th subtree, and `$w_i$` is the sample weight. This structure employs a residual-driven incremental learning strategy to progressively approximate the optimal classification boundary in each iteration. As this stage focuses on high-recall preliminary identification of highly suspicious samples, the model outputs multi-value risk level labels. These provide sample selection criteria and confidence distribution information for subsequent deep recognition layers.

3.3 Second Stage: Deep Learning-Based Fine Identification Layer

The second-stage precision identification layer undertakes the core task of extracting deep semantic and temporal correlation features from coarse-screened samples. It employs a hybrid architecture combining Convolutional Neural

Networks (CNN) and Long Short-Term Memory (LSTM) networks to achieve high-dimensional representation of transaction behavior sequences and dynamic risk identification. Model inputs include time-series transaction amounts, time intervals, geographic coordinates, and behavioral pattern vectors. After extracting local patterns through multiple convolutional layers, LSTM units capture long-term dependency features. The loss function for this stage comprehensively considers classification error and time-dependent penalty terms, specifically expressed as:

$$L_{deep} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log\left(\hat{y}_i\right) + \left(1-y_i\right)\log\left(1-\hat{y}_i\right)\right] + \alpha\sum_{t=1}^{T}\left\|\mathbf{h}_t - \mathbf{h}_{t-1}\right\|_2^2 \quad (3)$$

where `$N$` denotes the number of samples, `$y_i$` represents the true labels, `$\hat{y}_i$` indicates the predicted probability, `$\mathbf{h}_t$` is the hidden state vector of the LSTM at time step `$t$`, and `$\alpha$` is the smoothing constraint coefficient. The model iteratively updates parameters through backpropagation and the Adam optimizer, progressively achieving stable fitting of multidimensional temporal features. The high-confidence risk vectors output in this stage are fed into the third-stage ensemble decision layer for comprehensive final risk level determination, ensuring the system's recognition accuracy and generalization performance under multidimensional inputs [4].

3.4 Stage 3: Stacking-Based Ensemble Decision Layer

The ensemble decision layer employs the Stacking framework for cross-model evidence fusion. The base learners comprise GBDT and CNN-LSTM, while the meta-learner utilizes temperature-scaled logistic regression to output calibratable risk scores. During training, this layer generates out-of-fold probabilities for each base learner via K-fold cross-validation as meta-features. It further combines prediction variance, Shannon entropy, and time decay weights into uncertainty features, forming the vector $\mathbf{z}_i$. The objective function for this layer is defined as:

$$L_{meta} = -\sum_{i}\omega_i\left[y_i \log\left(s_i\right) + \left(1-y_i\right)\log\left(1-s_i\right)\right] + \lambda\left\|\mathbf{w}\right\|_2^2 + \alpha\sum_{j,t}\left(s_{j,t} - s_{j,t-1}\right)^2 \quad (3)$$

where $s_i = \sigma\left(\mathbf{w}^{\mathrm{T}}\mathbf{z}_i + b\right)$ denotes the probability output by the meta-learner, y_i represents the true label, ω_i indicates the sample weight based on the time window, $\mathbf{w}$ is the parameter vector, λ is the L2 regularization coefficient, α is the temporal smoothing coefficient, and (j,t) denotes the account sequence and time step. This layer employs a threshold strategy with multi-level alerts based on adaptive quantile settings. The calibration process constrains the desired calibration error to ensure score reliability [5]. The layer's output simultaneously returns the base learner contribution and uncertainty metric, providing stable inputs for subsequent alert triggering and online updates.

4 DYNAMIC RISK EARLY WARNING AND ADAPTIVE UPDATE MECHANISM

4.1 Time-Series Modeling and Risk Scoring Function Design

The time-series modeling module in the dynamic risk early warning system takes multi-scale behavioral sequences as input. It constructs a risk scoring function framework by integrating state transition patterns and behavioral intensity functions to model the evolutionary trends of time-series behaviors. This module adopts a hierarchical architecture design. The base

layer extracts feature sequences of account operations within fixed time intervals through a sliding window mechanism. Behavioral types include transaction density, time intervals, frequency of geographic coordinate jumps, and abnormal authentication events, among others [6]. The intermediate layer constructs a nested Gated Recurrent Unit (GRU) structure to capture cross-temporal dependencies and short-term volatility features within behavioral sequences. Building upon this, the top layer defines the risk scoring function R_t , which quantifies the current time step through a weighted integration of prediction errors and historical state metrics. This function is defined as follows:

$$R_t = \sigma\left(\sum_{k=1}^{T} \gamma_k \times \left| \hat{y}_{t-k} - y_{t-k} \right| + \beta \times \sum_{k=1}^{T} e^{-\lambda k} \times S_{t-k} \right) \quad (4)$$

Among these, `$R_t$` denotes the risk score at time step `$t$`; `$\hat{y}_{t-k}$` represents the model prediction values for the preceding `$k$` time steps; `$y_{t-k}$` corresponds to the actual labels; `$\lambda_k$` is the prediction bias weight coefficient; `$\beta$` is the historical state contribution adjustment factor; `$S_{t-k}$` indicates the behavioral state score at time step `$t-k$`; `$\lambda$` is the temporal decay parameter; and `$\sigma(\cdot)$` is the Sigmoid mapping function used to normalize the score output. Table 1 details the temporal dimension partitioning and sampling frequency for each feature type in the input sequence, aiding in defining the temporal granularity settings for behavioral trajectory inputs.

Table 1 Input Features and Time Dimension Settings for the Temporal Modeling Module

Feature Type	Description	Time Granularity	Sliding Window Length
Transaction Density	Number of Transactions per Unit Time	1 minute	30 minutes
Geo-Jump Frequency	Cross-city operation frequency	10 minutes	1 hour
Identity Verification Gap	Time Gap Between Consecutive Verifications	1 hour	6 hours
Cross-device Switching	Cross-device switching tagging frequency	5 minutes	1 hour

4.2 Multi-level Risk Thresholds and Early Warning Trigger Mechanism

Within the dynamic financial risk identification system, the accuracy of the early warning mechanism relies on the precise mapping between risk scores and risk level thresholds. To address uncertainties in scoring intervals and sample distribution heterogeneity, the warning system employs a multi-level threshold setting method based on quantile adaptive adjustment. This combines dynamic confidence interval segmentation to define distinct risk level ranges. Specifically, the risk score sequence is defined as: $\{S_t\}_{t-1}^{T}$ The empirical distribution function is computed via a sliding window: $F_S(s)$ The q th percentile threshold is then determined as: θ_q i.e., $F_S(\theta_q) = q$ A risk activation function is introduced at the early warning strategy layer:

$$A_t = \sum_{q=1}^{Q} \mathrm{I}(S_t - \theta_q) \times w_q \quad (5)$$

where A_t denotes the risk activation score at time t , $\mathrm{I}(\cdot)$ is the indicator function, w_q is the quantile weight, and θ_q is the quantile threshold corresponding to the risk level. This mechanism enhances responsiveness to high-frequency disturbances and structural mutations by adjusting w_q and window length. To further quantify response outcomes, the system implements a three-tier risk response (P1, P2, P3) corresponding to 90%, 75%, and 50% percentile trigger thresholds. Upon activation, it triggers historical trajectory comparisons and behavioral model replay mechanisms to enable rapid intervention and risk control strategy deployment [7]. Figure 2 illustrates the dynamic adaptation curve of multi-level thresholds and corresponding response interval divisions, providing an intuitive understanding of the model's tolerance boundaries for score fluctuations across different time periods.

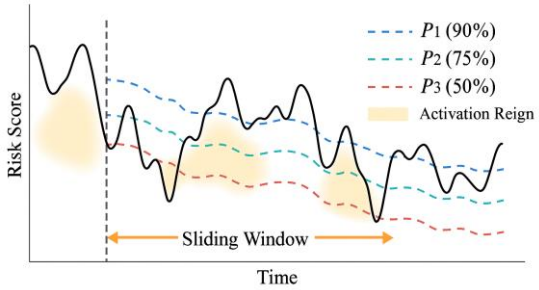

Figure 2 Multi-level Risk Score Threshold Adaptation Mechanism Diagram

4.3 Model Self-Learning and Online Update Strategy

To enhance the model's adaptability to data drift and behavioral pattern evolution in fintech environments, an auto-learning and online update mechanism based on error perception and confidence assessment is designed. This mechanism uses prediction results and feedback labels within a sliding time window as inputs to construct a multidimensional trigger function that determines whether the model requires local updates. The system defines the joint error function as follows:

$$L_t = \alpha \times \left| y_t - \hat{y}_t \right| + \beta \times \mathrm{KL}\left(P_t \middle\| P_{t-1} \right) + \gamma \times \left\| \theta_t - \theta_{t-1} \right\|_2 \quad (6)$$

where y_t is the true label at the current time step, $\hat{y}_t$ is the model's output probability, P_t and P_{t-1} represent the prediction distributions at the current and previous time steps, respectively, θ denotes the current model parameter vector, and α、β、γ is the hyperparameter weight controlling the contribution of different error terms to the total loss. When the error function L_t exceeds the dynamically set threshold δ_t , the system activates the online fine-tuning process. Model updates adopt a modular design, performing localized retraining only on network layers corresponding to sensitive feature paths to reduce computational burden and prevent degradation of global generalization capabilities [8]. The learning rate dynamically adjusts via exponential decay: $\eta_t = \eta_0 \times e^{-\lambda t}$, where λ is the decay coefficient, ensuring controllable convergence during the stable phase. For data collection, the self-learning module introduces confidence-guided sampling. Samples with confidence intervals exceeding preset bounds receive high-weight labeling and are prioritized in the online training queue. This builds a dynamically evolving self-labeled training set, supporting continuous updates to model architecture and strategies [9].

5 MODEL EXPERIMENTATION AND PERFORMANCE VALIDATION

5.1 Experimental Data Sources and Test Platform Construction

The experimental data originates from anonymized transaction logs and risk-labeled data provided by a domestic fintech platform. It encompasses approximately 11 million valid records spanning user behavior logs, device identifiers, transaction paths, authentication operations, and historical risk control outcomes from January 2023 to June 2024. The data has undergone cleaning and multidimensional label alignment, exhibiting typical characteristics of temporal sequence, anonymity, and class imbalance. The experimental platform was deployed on a high-performance workstation equipped with an NVIDIA RTX A6000 GPU, an Intel Xeon Gold 6330 CPU (2.0GHz, 56 cores), and 256GB of memory. The operating system was Ubuntu 22.04, and the deep learning environment was built upon Python 3.10, TensorFlow 2.12, and the LightGBM 3.3 framework [10]. Figure 3 illustrates the platform's system architecture, encompassing data preprocessing, feature engineering, model training, and online deployment modules.

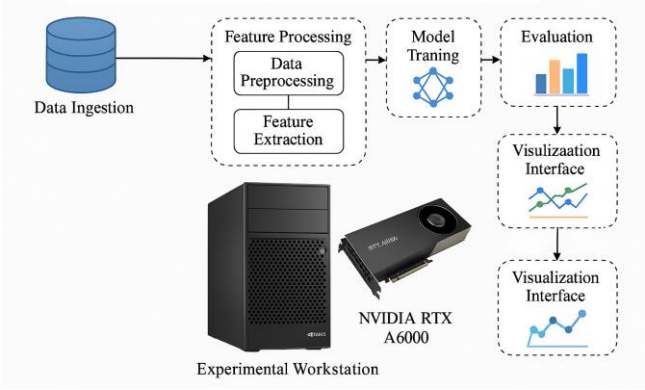

Figure 3 Experimental System Platform Architecture Diagram

5.2 Model Comparison Experiment Design

To validate the effectiveness and robustness of the proposed multi-stage integrated risk identification model on real fintech data, an evaluation experiment was designed incorporating five mainstream comparison models. These cover shallow decision tree algorithms, deep neural networks, sequence modeling networks, and hybrid architectures. The comparison models include GBDT, XGBoost, BiLSTM, CNN-LSTM, and Transformer. The first two models were used to validate structural feature modeling capabilities, while the latter three were employed to compare temporal modeling performance. As a multi-stage fusion architecture, the proposed model compares the stability and generalization capabilities of the aforementioned standalone models through a Stacking ensemble strategy and a dynamic risk scoring mechanism. All models were trained under a unified data preprocessing workflow and feature construction system. Hyperparameter tuning employed a grid search strategy, with the validation set proportion uniformly set at 20%. Table 2 details the model categories, input dimensions, key parameters, and output formats used in the experiments, facilitating subsequent result analysis and performance attribution.

Table 2 Model Configuration and Structural Parameter Comparison

Model Name	Architecture Type	Input Dimensions	Output Format	Key Features
GBDT	Tree Model	Structured Static Features	Risk Level Label	High Efficiency

XGBoost	Tree Model	Weighted Structural Features	Binary Classification Probability	Overfitting resistance
BiLSTM	Sequential Neural Network	Time Series Behavior Vector	Sequence Labels	Bidirectional Dependency
CNN-LSTM	Hybrid Neural Network	Multi-scale behavioral sequences	Probability Vectors	Local + Global Modeling
Transformer	Attention Model	Multimodal Input Sequences	Multi-class Scores	Long-Range Modeling Capability

5.3 Results Analysis and Interpretability Validation

To comprehensively evaluate the performance of the multi-stage ensemble model in fintech risk identification scenarios, this section compares the experimental results of the proposed model with five mainstream models across multiple metrics. Evaluation metrics include accuracy, recall, F1-score, AUC (area under the ROC curve), and KS value (Kolmogorov–Smirnov). Recall and AUC serve as key indicators for measuring the model's ability to identify high-risk samples. Table 3 results demonstrate that the Stacking ensemble model achieves the maximum envelope area across all metric dimensions, indicating its significant advantage in balancing recognition accuracy with distributional discrimination capability. Traditional tree-based models (e.g., GBDT, XGBoost) exhibit acceptable performance in Accuracy but fall significantly below deep structures in Recall and AUC—core metrics reflecting anomaly detection capability—revealing their insufficient modeling capacity for complex behavioral sequences. BiLSTM and CNN-LSTM, possessing time-dependent structures, demonstrated overall recognition performance markedly superior to single-structure models. Notably, CNN-LSTM exhibited relatively balanced performance on F1 and KS scores, indicating advantages in precise recognition and stability. Transformer models, benefiting from attention mechanisms, showed strong modeling capabilities for long sequences but exhibited slight fluctuations in scenarios with insufficient samples. Overall, ensemble models demonstrate superior metric balance and extreme risk capture, validating the adaptability of multi-stage architectures in complex financial trading scenarios.

Table 3 Performance Metrics Comparison Across Models on the Test Set

Model Name	Accuracy	Recall	F1-score	AUC	KS Value
GBDT	0.881	0.765	0.793	0.874	0.511
XGBoost	0.887	0.779	0.804	0.886	0.528
BiLSTM	0.901	0.823	0.842	0.902	0.561
CNN-LSTM	0.912	0.847	0.866	0.919	0.578
Transformer	0.908	0.834	0.853	0.912	0.569
Stacking Integrated Model	0.926	0.886	0.893	0.941	0.601

To validate model interpretability, SHAP (SHapley Additive exPlanations) analysis was further introduced to rank key feature contributions within the CNN-LSTM submodel. Figure 4 reveals that the top-ranked feature is "Transaction interval," whose high importance indicates that frequent or anomalous transaction intervals exert a core influence on the model's risk assessment. Next are "Cross-device switch frequency" and "Geo jump coefficient," representing device switching frequency and geographic jump patterns respectively. These features capture spatial anomalies in account activity, playing a critical role in detecting device spoofing and geographic drift attacks. "Transaction amount" and "Transaction time" rank mid-tier, indicating they serve as supplementary decision features within the model. Periodic behavioral features like "Day of week" and "Hour of day" also rank within the top ten, demonstrating that temporal patterns aid in identifying suspicious behavior. Lower-ranking features such as "Auth failures" and "Device model," while individually less impactful, may contribute significantly through higher-order nonlinear interactions when combined with

other features in modeling. Overall, the SHAP plots clearly reveal the model's decision transparency and logical consistency, providing feature explanations for subsequent risk control rule setting.

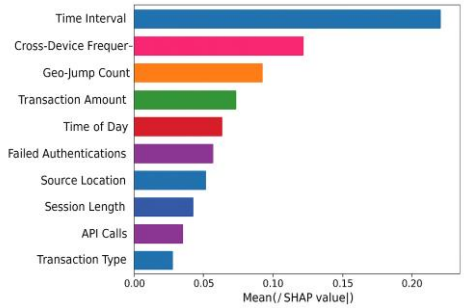

Figure 4 Key Feature Importance Ranking Based on SHAP

5.4 Model Stability and Robustness Evaluation

To validate the multi-stage ensemble model's adaptability to data fluctuations and structural perturbations in practical applications, this section evaluates model stability and robustness across three dimensions: temporal intervals, input noise perturbations, and category distribution imbalance. The temporal dimension test utilizes data spanning six quarters from Q1 2023 to Q2 2024 to assess performance fluctuations during long-term operation. Robustness testing against perturbations applied additive Gaussian noise to selected input variables at intensities of ±5%, ±10%, and ±15%. Imbalanced data testing simulated extreme scenarios of scarce high-risk samples by subsampling to control positive-to-negative sample ratios at 1:5, 1:10, and 1:20.

As shown in Table 4, our model maintains the highest AUC and F1-score across all test conditions while exhibiting the lowest standard deviation, demonstrating stable risk identification capabilities. For instance, under 1:10 class imbalance, the traditional GBDT model's AUC drops to 0.794, whereas our model remains at 0.889 with an F1-score of 0.765—significantly outperforming other models. In time-shift testing, our model exhibited the smallest fluctuation across quarterly data (Std. Dev. = 0.018), indicating excellent long-term generalization and robustness against variations, making it suitable for complex, dynamic financial environments.

Table 4 Robustness Evaluation Results of Models Under Different Perturbation and Imbalance Conditions

Test Scenario	Model	AUC	F1-score	Std. Dev.
Gaussian Perturbation ±5%	GBDT	0.842	0.771	0.029
	CNN-LSTM	0.873	0.794	0.022
	Model in this paper	0.917	0.835	0.015
Class imbalance 1:10	GBDT	0.794	0.682	0.044
	CNN-LSTM	0.826	0.711	0.036
	Model in this paper	0.889	0.765	0.021
Time Drift (Q1→Q4)	GBDT	0.807	0.715	0.038
	CNN-LSTM	0.855	0.742	0.027
	Model in this paper	0.903	0.789	0.018

6 CONCLUSION

This study constructs a multi-stage integrated recognition and dynamic early warning model for fintech crime risks, achieving joint modeling of structural features and temporal behaviors with online self-learning optimization. Experimental validation demonstrates the model's outstanding stability and robustness in complex environments, effectively supporting

real-time risk monitoring and precise early warning under high-dimensional data conditions. Future research may deepen the model's adaptability through cross-platform data fusion, heterogeneous feature transfer, and intelligent regulatory collaboration, providing a higher-level intelligent support system for financial security governance.

REFERENCES

[1] Ekundayo F, Atoyebi I, Soyele A, et al. Predictive analytics for cyber threat intelligence in fintech using big data and machine learning[J]. Int J Res Publ Rev, 2024, 5(11): 1-15.

[2] Metawa N, Metawa S. Internet financial risk early warning based on big data analysis[J]. American Journal of Business and Operations Research, 2021, 3(1): 48-60.

[3] Angela O, Atoyebi I, Soyele A, et al. Enhancing fraud detection and prevention in fintech: Big data and machine learning approaches[J]. World J. Adv. Res. Rev, 2024, 24(2): 2301-2319.

[4] Udeh E O, Amajuoyi P, Adeusi K B, et al. The role of big data in detecting and preventing financial fraud in digital transactions[J]. World Journal of Advanced Research and Reviews, 2024, 22(2): 1746-1760.

[5] Rahardja U, Miftah M, Rakhmansyah M, et al. Revolutionizing financial services with big data and fintech: A scalable approach to innovation[J]. ADI Journal on Recent Innovation, 2025, 6(2): 118-129.

[6] Zhou Z, Li L, Huang P. A Comprehensive Multi-Dimensional Risk Monitoring Model for Illegal Financial Activities[J]. Journal of Risk Analysis and Crisis Response, 2024, 14(4): 472-488

[7] Popoola N T. Big data-driven financial fraud detection and anomaly detection systems for regulatory compliance and market stability[J]. Int. J. Comput. Appl. Technol. Res, 2023, 12(09): 32-46.

[8] Ekundayo F. Big data and machine learning in digital forensics: Predictive technology for proactive crime prevention,'[J]. Complexity, 2024, 24(2): 2692-2709.

[9] Hasan M, Hoque A, Le T. Big data-driven banking operations: Opportunities, challenges, and data security perspectives[J]. FinTech, 2023, 2(3): 484-509.

[10] Fatunmbi T O. Advanced frameworks for fraud detection leveraging quantum machine learning and data science in fintech ecosystems[J]. World Journal of Advanced Engineering Technology and Sciences, 2024, 12(01): 495-513.

Project Name：Host of the National Social Science Fund project "Research on the Theory, Policy and Prospects of Internet Finance" (project number: 16BJY178), From June 2016 to November 2021