Article

Not peer-reviewed version

A Micro-Manifold Identity-Preserving Spatiotemporal Graph Neural Network for Financial Risk Early Warning

Jin Kuang , Fusheng Chen , Te Guo , Chiawei Chu *

Posted Date: 24 March 2026

doi: 10.20944/preprints202603.1902.v1

Keywords: financial risk early warning; spatiotemporal graph neural network; over-smoothing; extreme imbalanced learning; explainable artificial intelligence

Article

A Micro-Manifold Identity-Preserving Spatiotemporal Graph Neural Network for Financial Risk Early Warning

Jin Kuang [1], Fusheng Chen [2], Te Guo [2] and Chiawei Chu [2,*]

[1] School of International Business (The Belt and Road School of Foreign-Related Rule of Law), Zhejiang International Studies University, Hangzhou, China
[2] Faculty of Data Science, City University of Macau, Macau, China
* Correspondence: cwchu@cityu.edu.mo

Abstract

Traditional financial early warning models often rely on the independent and identically distributed (IID) assumption, failing to adequately capture cross-sectional spatial contagion effects and temporal dynamic mutations, and are susceptible to the over-smoothing problem when processing highly imbalanced graph networks. To address these limitations, this study proposes a micro-manifold-based identity-preserving spatiotemporal graph neural network framework (Micro-STAGNN). In the spatial dimension, an identity-preserving graph convolutional operator (IP-GCN) is constructed. By hard-coding a self-preservation coefficient ($\lambda = 0.8$), it quantifies peer risk spillover while mitigating feature dilution, ensuring the transmission of heterogeneous default signals. In the temporal dimension, Long Short-Term Memory networks are cascaded with a temporal attention mechanism to capture the nonlinear temporal inflection points that trigger financial distress. The empirical study utilizes a sample of China's A-share market from 2015 to 2025, evaluating the model using an Out-of-Time Validation protocol and Focal Loss. Results indicate that under a highly imbalanced distribution with a positive-to-negative sample ratio of approximately 1:50, Micro-STAGNN achieves an OOT ROC-AUC of 0.9095, a minority class default recall of 89%, and reduces the missed detection rate to 11%, outperforming traditional nonlinear cross-sectional models such as XGBoost. Furthermore, temporal attention weights provide explainable support for the early warning results.

Keywords: financial risk early warning; spatiotemporal graph neural network; over-smoothing; extreme imbalanced learning; explainable artificial intelligence

MSC: 68T07; 91G40; 62P05; 91G15

---

1. Introduction

In the new generation of information technology environment, the financial risk landscape faced by listed companies is becoming increasingly complex, extending from the deterioration of single financial indicators to multidimensional compounding trends such as spatial risk contagion and cross-boundary risk transmission. This complexity drastically exacerbates the uncertainty in measuring and integrating risk evaluation information, making it difficult for traditional single-risk scenario research to comprehensively capture the key factors affecting a firm's financial health [1–3]. Consequently, financial risk early warning has become increasingly critical as a key mechanism for effectively managing risks and promoting sustainable corporate development.

Financial risk early warning for listed companies aims to deeply deconstruct historical operational trajectories, thereby providing quantified decision support for capital market stakeholders to take proactive preventive measures. However, traditional early warning methods are often confined to treating enterprises as isolated, independent and identically distributed (IID)

 doi:10.20944/preprints202603.1902.v1

samples, lacking in-depth exploration of implicit correlations and risk spillover effects within peer competition chains. Furthermore, these comprehensive evaluations fail to overcome the temporal lag caused by the Markov smoothing assumption. In the face of intricate enterprise correlation networks in real capital markets, thoroughly mining cross-sectional spatial topologies and temporal nonlinear mutation features has become a crucial pathway to enhancing early warning accuracy [4]. Concurrently, the over-smoothing problem in graph networks caused by high-dimensional spatiotemporal feature fusion, along with extreme sample imbalance under long-tail distributions, presents immense algorithmic challenges for current deep learning-based financial risk early warning systems.

When listed companies fail to effectively absorb external shocks and internal operational pressures, their normal operational activities suffer direct, severe impacts, inevitably leading to financial deterioration and a significantly higher probability of bankruptcy default. The financial risk of listed companies is not an isolated microeconomic phenomenon; its negative externalities propagate through the capital market network to various stakeholders, generating extensive spatial spillover consequences. Specifically: for the company itself, financial risk not only erodes its market value but fundamentally constrains its long-term prospects; for investors, the deterioration of underlying assets is tightly linked to capital security, and the outbreak of a financial crisis directly results in severe capital erosion; for regulatory bodies, if individual default events accumulate and diffuse within the topological network, they directly threaten the overall risk-resistance capacity of the macroeconomic financial system [5,6]. Thus, the financial risk of listed companies transcends micro-level management and has become a core macroeconomic issue involving capital market stability and systemic risk prevention.

To address the structural deficiencies of traditional financial early warning models in extracting cross-sectional risk spillovers and temporal dynamic evolution features, this study proposes a micro-manifold-based identity-preserving spatiotemporal graph neural network framework (Micro-STAGNN). To circumvent intertemporal concept drift triggered by macroeconomic cyclical fluctuations, this architecture strictly strips away global macroeconomic noise, focusing exclusively on deconstructing the spatiotemporal dynamics of endogenous corporate financial indicators [7,8]. In the spatial dimension, this study constructs an Identity-Preserving Graph Convolutional Network (IP-GCN). While capturing the spatial contagion effects of systemic risk within peer competition chains, this operator utilizes a hard-coded feature preservation coefficient (i.e., forcibly anchoring 80% of individual micro-features) to algebraically block the feature over-smoothing collapse prone to occur when processing highly imbalanced graph topologies, ensuring the complete transmission of heterogeneous risk signals from defaulting enterprises. In the temporal dimension, the model cascades Long Short-Term Memory (LSTM) networks with Temporal Attention to break through the smoothing assumption of traditional Markov chains. This mechanism adaptively performs targeted weighting on historical hidden states, precisely locking onto and amplifying nonlinear mutation inflection points that trigger financial bankruptcy (such as anomalous high leverage or liquidity depletion signals). This not only effectively blocks the gradient decay of critical early warnings but also provides clear economic attribution and explainability for the deep learning early warning system via attention weight heatmaps.

Regarding optimization objectives and empirical protocols, addressing the authentic long-tail distribution of ST (Special Treatment) default samples in the A-share market (positive-to-negative ratio of approx. 1:50), the model introduces Focal Loss to asymmetrically reshape the decision hyperplane, maximizing the recall baseline for extreme risks. Furthermore, this study abandons traditional static randomized cross-validation in favor of an absolutely isolated Out-of-Time Validation protocol. By utilizing historical graph snapshots from 2018-2022 for network optimization and executing strict extrapolative predictions in the future time window of 2023-2025, this study quantitatively confirms the model's generalization capability and missed-detection prevention efficacy in real-world systemic financial risk prevention scenarios. The contributions of this paper are as follows:

 doi:10.20944/preprints202603.1902.v1

At the data manifold level, an early warning foundation of purely micro-endogenous financial features is constructed. By strictly stripping away global macroeconomic cyclical noise, the interference of intertemporal concept drift is eliminated, establishing an evolutionary risk measurement standard driven by endogenous corporate operational trajectories.

At the spatial topology level, an Identity-Preserving Graph Convolutional Network (IP-GCN) is designed. While quantifying industry risk spillover effects, this operator algebraically blocks the severe feature over-smoothing prevalent under long-tail distributions, ensuring the complete transmission of heterogeneous default signals across spatial links.

At the temporal evolution level, LSTM and temporal attention mechanisms are cascaded to break the Markov smoothing assumption. This mechanism not only prevents gradient decay of long-sequence early warnings and precisely targets nonlinear risk mutation inflection points, but also provides white-box economic attribution through weight heatmap outputs, enhancing the explainability of the deep learning early warning framework.

2. Literature Review

Since the 1930s, foreign scholars first proposed the concept of “financial crisis” and conducted exploratory research on this corporate phenomenon. As the issue spread, Fitzpatrick boldly hypothesized as early as 1932 that the formation of a financial crisis is a gradual, cumulative process that can be predicted. Subsequently, numerous scholars began researching financial crisis prediction, utilizing various predictive methods and statistical tools to conduct extensive experiments. By comparatively analyzing experimental results, they selected prediction models with the highest accuracy to achieve financial early warning effects [9]. Financial risk early warning models have evolved from traditional statistical models to machine learning models, and further to deep learning models.

2.1. Research on Financial Early Warning Based on Statistical Models

Early research on financial crisis prediction relied primarily on statistical methods, with representative models including the Altman Z-score and logistic regression models. The former constructs a linear combination of financial indicators through discriminant analysis, while the latter estimates the probability of financial distress based on specific corporate features [10–12]. Although these models provided foundational reference value in corporate risk assessment, their effectiveness is constrained by methodological assumptions.

Typically, when selecting sample companies for financial early warning research, a pairing principle is followed, meaning an equal number of ST and non-ST companies are selected, or the criteria for determining if a sample is in financial distress are matched. Chen Xiao and Chen Zhihong (2000) used 132 healthy companies and 38 ST companies as research subjects, applying logistic regression to predict financial crises of Chinese listed companies in an attempt to build an effective financial early warning model [13]. Considering the potential correlations among factors triggering financial crises, Wang Lin and Zhou Xin (2007) selected 14 financial indicators based on the four-ability theory of financial analysis as influencing factors for their research. Simultaneously, they sampled four consecutive years of financial data from 110 listed companies (including 55 ST companies) across the Shanghai and Shenzhen exchanges, utilizing principal component analysis for predictive research [14].

2.2. Research on Financial Early Warning Based on Machine Learning

Traditional statistical models rely heavily on strict data distribution assumptions and linear prerequisites, prompting researchers to shift toward more powerful data-driven methods. Compared to traditional predictive models, machine learning models do not require specific prior knowledge assumptions, leading to their gradual application in financial early warning research by domestic and international scholars. Machine learning is a discipline that learns patterns from data for

 doi:10.20944/preprints202603.1902.v1

prediction and analysis; its core lies in modeling data mathematically and optimizing parameters algorithmically to fulfill learning objectives. Specifically, models such as Naive Bayes (NB), Support Vector Machines (SVM), Decision Trees (DT), Artificial Neural Networks (ANN), and Random Forests (RF) have been widely applied [15–18]. Danenas et al. utilized SVM models to identify listed companies with default risk, finding that linear SVM classifiers offered faster training times and higher accuracy [19]. Kazemi et al. proposed using the NB model to estimate financial risk probabilities, verifying its effectiveness through evaluations on real listed company datasets [20]. Ciampi et al. introduced ANN models into financial risk early warning for listed companies, discovering that ANNs could enhance the early warning sensitivity of models like LR, and performed exceptionally well for smaller companies [21]. Uddin applied the RF method to financial early warning, proving its effectiveness while noting that RF could provide stronger explainability compared to other classifiers [22].

2.3. Application Research Based on Deep Spatiotemporal Features and Graph Networks

Deep learning, with its powerful nonlinear mapping capabilities, has gradually become the frontier paradigm for financial crisis early warning. Early deep learning applications focused primarily on local feature extraction and temporal dependencies. For instance, Wang Xuefeng (2023) used Convolutional Neural Networks (CNN) to extract salient local features from financial indicators; Barthelemy et al. (2024) demonstrated that LSTM possesses superior sequential capture capabilities compared to logistic regression when predicting currency crises [23,24]. However, these classic architectures treat enterprises merely as isolated nodes, entirely ignoring peer competition correlations and cross-sectional risk spillovers in real capital markets.

To address the absence of spatial topology, Graph Convolutional Networks (GCN) and their variants were introduced for financial fraud and default detection. Wang (2024) proposed a multi-relational graph convolutional network (FraudGCN), combining focal loss to handle extreme fraud sample imbalance [25]; the KeGCN_R model by Wang et al. (2025) utilized knowledge enhancement to improve graph robustness [26]; Yang (2025) constructed a heterogeneous graph convolutional network to identify digital financial fraud risks [27]. These studies confirmed the effectiveness of graph structures in capturing systemic risk contagion.

Recently, some scholars have attempted to jointly utilize temporal and spatial domain operators. Zhang (2025) simultaneously captured enterprise correlations and dynamic evolution features by fusing GNN, GAT, and LSTM [28]; Liu et al. (2025) proposed a framework integrating self-attention mechanisms with GCN to analyze multidimensional features and dynamic transaction correlations of financial entities [29]. The aforementioned studies provide important theoretical support for financial early warning. However, when confronting the authentic extremely imbalanced (approx. 1:50) long-tail data of the A-share market, the fusion of high-dimensional spatiotemporal features easily triggers graph over-smoothing, causing sparse heterogeneous default signals to be completely diluted by the majority of healthy nodes. Addressing this core pain point, this study innovatively proposes the Micro-feature Spatiotemporal Attention Graph Neural Network framework (Micro-STAGNN). By strictly defending the heterogeneous baseline cross-sectionally via the strong Identity-Preserving operator (IP-GCN) and synergizing with the temporal attention mechanism to precisely target deterioration inflection points, this model achieves absolute defense against long-tail tail risks under an extrapolative protocol that prevents future data leakage.

In summary, based on existing domestic and international research, the development of financial risk early warning models exhibits an evolution from static prediction to dynamic temporal evolution, and from black-box models to explainability enhancement. These evolutions not only drive continuous methodological innovation in financial early warning research but also provide critical theoretical support and technical reference for the model construction of this study [30–32]. Addressing the limitations of traditional models in capturing cross-boundary risk contagion and dynamic temporal features, this study innovatively proposes the Micro-STAGNN framework. This model captures systemic risk contagion across industries while preserving enterprise heterogeneity via a residual graph convolutional network, and simultaneously combines temporal attention mechanisms to precisely profile the dynamic deterioration trajectories of financial health, significantly enhancing the explainability of the early warnings. Experiments utilizing strict cross-temporal graph snapshot extrapolative validation demonstrate that the model exhibits exceptional robustness and proactive predictive capability in addressing sample imbalance and preventing future data leakage.

3. Methodology

Figure 1 presents the end-to-end Computational Graph topology of the proposed Micro-STAGNN. This architecture aims to deconstruct the spatiotemporal dynamics of financial crisis evolution under extreme imbalanced distributions. The overall forward propagation link can be physically isolated into three cascaded mapping domains: first, endogenous financial feature tensors are mapped to a higher-order latent space via linear projection and layer normalization; subsequently, the core Identity-Preserving Graph Convolutional operator (IP-GCN) executes risk contagion calculations within the local industry homogeneity network across the cross-sectional dimension; finally, the intertemporal temporal manifold is synergized via LSTM and the temporal attention mechanism to precisely target nonlinear mutation inflection points in historical trajectories, asymmetrically reconstructing the final decision hyperplane through Focal Loss. The following subsections will sequentially conduct strict algebraic derivations for each physical module.

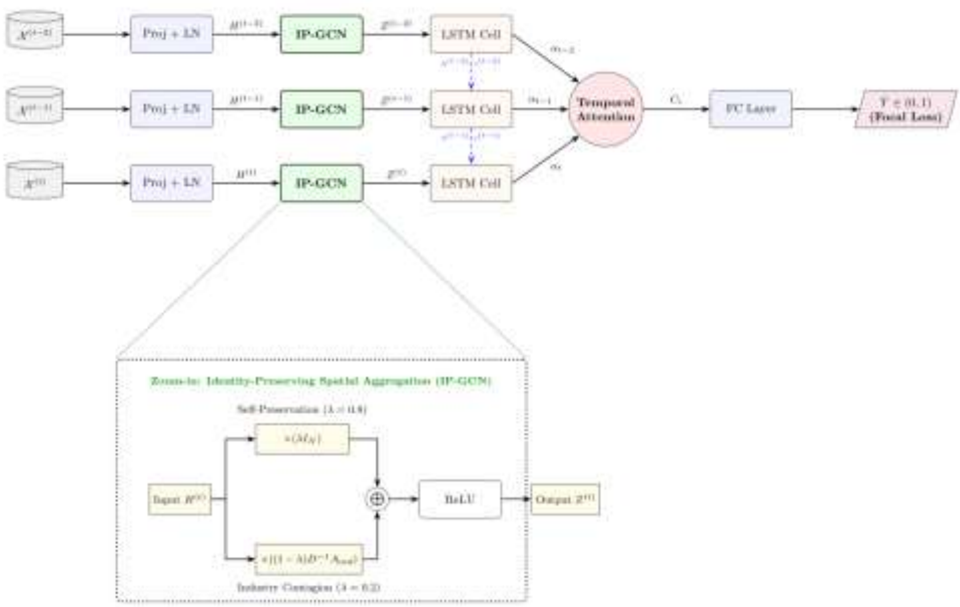

Figure 1. End-to-end computational graph topology of the proposed Micro-STAGNN.

3.1. Micro-Feature Manifold and Projection

In the financial network $G = (V, \mathcal{E})$ characterized by extreme imbalance and intertemporal regime drift, this study discards macroeconomic factors that easily induce signal-to-noise ratio collapse, constructing purely micro-endogenous financial feature tensors $X \in \mathbb{R}^{N \times T \times F}$. To stabilize feature distribution and accelerate convergence, the model executes linear projection and Layer Normalization on the input slice $X^{(t)} \in \mathbb{R}^{N \times F}$:

$$H^{(t)} = \mathrm{LayerNorm}(X^{(t)} W_{proj} + b_{proj}) \tag{1}$$

Where $W_{proj} \in \mathbb{R}^{F \times d}$ maps the 8-dimensional original indicators to a 64-dimensional high-order latent space, establishing the fundamental spatiotemporal representation $H^{(t)}$ of the nodes.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2026 doi:10.20944/preprints202603.1902.v1

3.2. Identity-Preserving Spatial Aggregation

When traditional Graph Convolutional Networks (GCN) process highly imbalanced (approx. 1:50) assortative graphs, they easily trigger Feature Dilution, causing the anomalous signals of a minority of defaulting entities to be masked by the homogenized representations of the majority of healthy nodes. To counter this flaw, Micro-STAGNN constructs an Identity-Preserving topological operator.

It assumes that the observed network entity correlations are driven by industry homogeneity. Specifically, at each observation cross-section (time snapshot), the model breaks the isolation walls of independent samples, physically merging all extant listed companies in the entire market into a single global graph network for joint topological inference. In this global graph, the edge establishment rules between nodes are as follows: for any nodes $v_i, v_j \in V$, the elements of their industry adjacency matrix $A_{ind} \in \mathbb{R}^{N \times N}$ obey the indicator function:

$$(A_{ind})_{ij} = \mathbb{I}(\mathrm{Ind}(v_i) = \mathrm{Ind}(v_j)) \tag{2}$$

Where $\mathrm{Ind}(\cdot)$ is a mapping function outputting the node's industry attribute. The normalized spatial aggregation weight matrix A_{norm} is defined as:

$$A_{norm} = \lambda I_N + (1-\lambda)D^{-1}A_{ind} \tag{3}$$

Where I_N is the identity matrix, and $D_{ii} = \sum_j (A_{ind})_{ij}$ is the degree matrix. The hyperparameter $\lambda \in [0,1]$ controls the aggregation ratio between individual heterogeneous features and peer spatial spillover features. Based on cross-validation from the historical training set (2018-2022), this architecture establishes $\lambda = 0.8$ as the empirical optimal solution, meaning it forcibly retains 80% of the target node's own micro-features, absorbing only 20% of peer risk diffusion. The graph spatial smoothing operation at time t is expressed as:

$$Z^{(t)} = \mathrm{ReLU}(A_{norm}H^{(t)}) \tag{4}$$

While quantifying the peer risk contagion links, this operator algebraically blocks the feature dilution phenomenon from the ground up, establishing the recall baseline for extreme risk samples.

3.3. Mutation-Aware Temporal Attention

After obtaining the spatially smoothed tensor sequence $Z = [Z^{(1)}, Z^{(2)}, \dots, Z^{(T)}]$, the model applies Long Short-Term Memory networks (LSTM) to extract the intertemporal dynamic trajectories of the balance sheets:

$$h_{i,t}, c_{i,t} = \mathrm{LSTM}(Z_i^{(t)}, h_{i,t-1}, c_{i,t-1}) \tag{4}$$

Financial crisis evolution inherently possesses nonlinear mutation characteristics. To break the Markov smoothing assumption of terminal states in long sequences, the model cascades a Temporal Attention mechanism to adaptively reconstruct historical states:

$$u_{i,t} = \tanh(W_a h_{i,t} + b_a) \tag{5}$$

$$\alpha_{i,t} = \frac{\exp(u_{i,t}^{T} v_a)}{\sum_{\tau=1}^{T} \exp(u_{i,\tau}^{T} v_a)} \tag{6}$$

$$C_i = \mathrm{LayerNorm}\left(\sum_{t=1}^{T} \alpha_{i,t} h_{i,t}\right) \tag{7}$$

The attention weight $\alpha_{i,t}$ enables the network to span the temporal domain, precisely locking onto early anomalous inflection points that trigger bankruptcy. The final context vector C_i is mapped to a scalar logit value via fully connected layers and Dropout operations, outputting the predicted default probability p_i.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2026 doi:10.20944/preprints202603.1902.v1

3.4. Asymmetric Optimization via Focal Loss

Addressing the decision hyperplane shift caused by the long-tail distribution in the authentic A-share context, the terminal end of this architecture discards standard binary cross-entropy, employing Focal Loss for asymmetric boundary reshaping:

$$\mathcal{L}_{focal} = -\frac{1}{N}\sum_{i=1}^{N} \alpha_i \, (1 - p_i)^{\gamma} \log(p_i) \tag{8}$$

Based on empirical grid optimization, this study fixes the focusing parameter $\gamma = 1.5$ and the default sample weight $\alpha = 0.75$. This cost-sensitive constraint, combined with dynamic Youden Index truncation, forcibly widens the manifold gap between the healthy majority class and the defaulting minority class, achieving a global Pareto optimum.

3.5. Computational Complexity Analysis

In real-time early warning practices within actual capital markets, algorithmic scalability is crucial. Let N be the total number of nodes, T be the length of the temporal window, d be the latent tensor dimension, and k be the average number of connected nodes within an industry. The time complexity of this architecture primarily consists of the graph aggregation module and the temporal module. For the spatial aggregation operation, since A_{ind} exhibits strict industry block-diagonal features, the time complexity of its sparse matrix multiplication collapses from $O(N^2 d)$ to $O(Nkd)$. The LSTM and temporal attention mechanisms execute mappings on a sequence of length T, bearing a time complexity of $O(NTd^2)$. Therefore, the overall forward propagation time complexity of Micro-STAGNN is:

$$O(Nkd + NTd^2)$$

Because T (constant 3) and d (constant 64) are far smaller than N, and $k \ll N$, the time complexity of this architecture maintains a strictly linear relationship with the market node scale N, fully satisfying the computational constraints for real-time risk monitoring of millions of entities.

4. Experiment Results

4.1. Data Description

This study selects all enterprises listed on domestic stock exchanges in China (including SSE, SZSE, and BSE) from 2015 to 2025 as research subjects. This sample interval covers 19 major industry categories, including agriculture, mining, and manufacturing, encompassing 5,724 surviving or listed enterprises, demonstrating excellent representativeness and breadth in both time span and industry coverage. Regarding financial capability evaluation, this paper integrates the core logic of DuPont analysis with traditional financial capability analysis dimensions, constructing an evaluation indicator system from four key dimensions: solvency, operational capability, profitability, and growth capability. Based on whether a risk warning has been implemented by the exchange, sample enterprises are classified into normal listed companies, ST companies, and *ST companies. Relevant financial data are sourced from the public financial statements of the sample companies to ensure objectivity and verifiability [9]. The specific initial sample variables are detailed in Table 1.

Table 1. Indicator system based on financial capability analysis.

Dimension	Specific Indicators
Solvency	Quick Ratio
	Asset-Liability Ratio
	Cash Flow Ratio
Operational Capability	Total Asset Turnover
	Accounts Receivable Turnover

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2026 doi:10.20944/preprints202603.1902.v1

	Inventory Turnover
Profitability	Return on Equity (ROE) Return on Assets (ROA) Operating Profit Margin
Growth Capability	Operating Revenue Growth Rate Net Profit Growth Rate

Data preprocessing is a critical prerequisite for building an effective financial early warning system. Given the presence of missing values, duplicate records, and unstructured issues in the raw data, this study processes them according to the following steps: First, samples with severe missing values in key financial indicators are eliminated, and imputation methods are used to fill minor missing values in specific indicators; Second, to prevent multicollinearity from interfering with model estimation, collinearity diagnostics are performed on initially selected indicators within the same dimension, removing redundant variables; Finally, to eliminate dimensional impacts, all retained indicators undergo standardization (normalization) processing. Through this screening, eight indicators are ultimately confirmed to construct the financial early warning system: Quick Ratio, Asset-Liability Ratio, Accounts Receivable Turnover, Inventory Turnover, Total Asset Turnover, Return on Assets (ROA), Return on Equity (ROE), and Operating Profit Margin.

To verify the separability of early warning features, this study performed grouped descriptive statistics and multicollinearity diagnostics on the 8 established financial indicators. As shown in Table 2, defaulting entities (ST group) and healthy entities exhibit significant statistical differences across core financial dimensions. Specifically, the mean Asset-Liability Ratio for the ST group is 0.5961 (higher than the healthy group's 0.4095), and the mean Quick Ratio is 1.5949 (lower than the healthy group's 2.2224), reflecting the debt pressure and liquidity characteristics of defaulting enterprises in the pre-crisis phase. In the profitability dimension, both ROA and ROE for the ST group present negative values (-0.0585 and -0.6492, respectively), demonstrating the capital erosion effect during financial distress.

Table 2. Descriptive statistics of features (derived from text context).

Financial Indicator	**Healthy Samples (is_ST = 0)**		**Default Samples (is_ST = 1)**	
	Mean	Std	Mean	Std
Quick Ratio	2.2224	2.958	1.5949	2.786
Asset-Liability Ratio	0.4095	0.2025	0.5961	0.2904
Accounts Receivable Turnover	1440.65	231413.2	44.7921	518.82
Inventory Turnover	322.3	21550.36	279.67	7149.7
Total Asset Turnover	0.6667	0.5316	0.4695	0.5386
ROA	0.0477	0.0982	-0.0585	0.1755
ROE	0.0608	0.5819	-0.6492	6.3412
Operating Profit Margin	-1.3484	244.85	-0.697	3.8368

Furthermore, this study executed Pearson Correlation and Variance Inflation Factor (VIF) tests. Test results indicate that the absolute values of correlation coefficients among all indicators in the feature matrix are below 0.6 (the highest negative correlation is -0.556 between Quick Ratio and Asset-Liability Ratio). Concurrently, all feature VIF values are distributed between 1.000 and 2.621, well below the conventional threshold for severe multicollinearity $VIF = 5$. These measurement results demonstrate that severe multicollinearity does not exist among the input features, satisfying the model's fundamental requirement for feature independence.

 doi:10.20944/preprints202603.1902.v1

4.2. Out-of-Time Validation Protocol

Traditional financial early warning research often employs static random cross-validation, a mechanism that easily induces Look-ahead Bias in temporal graph networks by disrupting the timeline. To authentically simulate forward-looking risk control scenarios, this study deploys an absolutely isolated intertemporal graph snapshot Out-of-Time Validation protocol.

Based on the full-market sample database established in Section 4.1, this architecture physically slices the data manifold along the time axis: taking t as the baseline observation point, the model retrospectively extracts continuous micro-feature tensors $\mathrm{X}^{(t-2:t)}$ of length $\mathrm{T}=3$ and the concurrent static industry topology $\mathrm{A}_{\mathrm{ind}}$ to predict the default status at period $t+1$. Given the configuration of the temporal sliding window length $\mathrm{T}=3$, the full sample data from 2015 to 2017 is strictly withheld as a Lookback Initialization Window for features, ensuring that the input tensors of the first batch of training targets possess complete intertemporal trajectories. Regarding dataset partitioning, this study designates the graph snapshot sequence from 2018 to 2022 as the historical optimization manifold (Training Set), while the Out-of-Time Test Set is strictly isolated within the future time window of 2023 to 2025. This protocol fundamentally eradicates survivorship bias from an algebraic level.

4.3. Evaluation Metrics and Baselines

Given the extremely imbalanced long-tail characteristics (approx. 1:50 positive-to-negative ratio) of the target observation network, traditional Accuracy exhibits severe statistical deception. Therefore, this study deploys high-order Threshold-independent Metrics and a cost-sensitive evaluation system. Core evaluation metrics include:

ROC-AUC: Serves as the global baseline, quantifying the classifier's expected probability of distinguishing between defaulting and healthy entities across all truncation thresholds. Optimal Threshold Metrics: Adaptively optimizes the Youden Index ($J = TPR - FPR$) via the ROC curve to establish the optimal dynamic truncation boundary, and outputs a complete classification report at this threshold covering Accuracy, Precision, Recall, and F1-Score. To establish the performance boundaries of the proposed Micro-STAGNN, this study deployed two categories of baseline models and ablation variants for defense testing under equivalent OOT protocols: Cross-sectional Detection Baselines: Logistic Regression (LR) acts as the classic IID linear detector, with forced class weight balancing injected during training; Random Forest (RF) sets 100 decision trees to benchmark the nonlinear ensemble limits of traditional machine learning on cross-sectional feature spaces. Network Ablation Variants: A degraded network stripped of the Identity-Preserving graph convolution module (w/o GCN), and a degraded network stripped of the temporal attention mechanism (w/o Attention), used to strictly quantify the independent algebraic contributions of the spatiotemporal dual-domain operators.

4.4. Implementation Details

This architecture completes end-to-end construction within the PyTorch deep learning framework, with the full model executing parallel tensor operations on an NVIDIA GeForce RTX 4070, and the underlying architecture developed on Python 3.9 and PyTorch 2.0. The high-order latent space dimension in model forward propagation is set to $d = 64$. At the optimizer level, the native Adam operator is used, with an initial learning rate set to $\eta = 0.002$, applying a Weight Decay of 1×10^{-4} to suppress overfitting in the parameter space.

For graph topology input features, this model discards traditional Mini-batch sampling, which destroys connectivity, and deploys Full-graph Snapshot Training. That is, in any given observation year (Epoch internal iteration), the model inputs all corporate nodes belonging to the same sector as an indivisible holistic graph into the network simultaneously. This mechanism ensures that the complete industry cross-sectional contagion links are losslessly preserved during forward propagation and risk prediction phases. The network executes 150 Epochs of global iterations.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2026 doi:10.20944/preprints202603.1902.v1

Regarding loss function configuration, the cost-sensitive constraints are strictly aligned with the authentic long-tail distribution, establishing Focal Loss's focusing parameter $\gamma = 1.5$ and default class weight $\alpha = 0.75$.

4.5. Out-of-Time Generalization

On the strict Out-of-Time (OOT) test set from 2023-2025, this study measured the actual generalization efficacy of each architecture when dealing with extremely imbalanced (1:50) financial manifolds and macroeconomic intertemporal drift. Table 3 (OOT Results) reports the global and minority class evaluation metrics under the adaptively optimal truncation threshold.

Table 3. Model performance evaluation.

Models	Accuracy	Precision	Recall	F1-Score	ROC-AUC
LR	0.8196	0.06	0.63	0.12	0.795
RF	0.8578	0.08	0.63	0.14	0.8193
XGBoost	0.8337	0.07	0.66	0.13	0.8296
RNN (Duan&Ren,2026)	0.9607	0.2092	0.402	0.2752	0.9014
Micro-STAGNN	0.817	0.08	0.89	0.14	0.9095

As shown in Table 3, empirical results establish the physical boundaries of different architectures in long-tail risk capture. LR, based purely on linear cross-sectional assumptions, presents a significant representational bottleneck, with an OOT ROC-AUC of only 0.7950, and its Recall in the sparse default sample interval drops to 0.63. This result quantitatively confirms the failure of the static IID assumption in extreme risk measurement within real capital markets. Although RF, which introduces nonlinear decision trees, raises the global ROC-AUC to 0.8193, its actual recall for defaulting entities stagnates at 0.63. Furthermore, the XGBoost algorithm, equipped with extreme nonlinear ensembling and severe cost-sensitive constraints (scale_pos_weight = 45.08), establishes the measurement ceiling for cross-sectional feature spaces after exhaustive parameter optimization, achieving an ROC-AUC of 0.8296. However, limited by the lack of physical dimensions for temporal evolution and spatial graph topologies, XGBoost's recall for defaulting entities still encounters an absolute upper bound blockage at 0.66. This series of baseline confrontations algebraically proves that merely stacking nonlinear parameters reliant on individual cross-sectional features fundamentally cannot break through the observational blind spots of intertemporal risk evolution and peer contagion effects.

Further comparing the time-series baseline model, although RNN captures the intertemporal evolutionary trajectories of enterprises through a hidden state recurrence mechanism [8], achieving an OOT ROC-AUC of 0.9014 (superior to traditional static cross-sectional models), this single-domain architecture exhibits significant limitations in adapting to highly imbalanced long-tail manifolds under standard truncation thresholds, with a minority class Recall of only 0.4020. The reason for this phenomenon is that standard RNN models tend to fit the majority class in exchange for high global Accuracy (reaching 0.9607), resulting in nearly 60% of true default risks being missed by the model. In contrast, Micro-STAGNN, relying on spatiotemporal dual-domain operators (IP-GCN combined with Temporal Attention), quantitatively captures systemic risk contagion across cross-sections while aligning with historical temporal evolution; paired with Focal Loss and adaptive threshold optimization, the model elevates the recall rate of default samples to 0.89 while securing the OOT ROC-AUC (0.9095), effectively controlling systemic missed detection risks, thereby algebraically achieving performance optimization over both the single temporal model (RNN) and the single cross-sectional model (XGBoost). Comparatively, the proposed Micro-STAGNN architecture establishes an overwhelming advantage on the global predictive probability distribution, with its OOT ROC-AUC leaping to 0.9095 (see Figure 2).

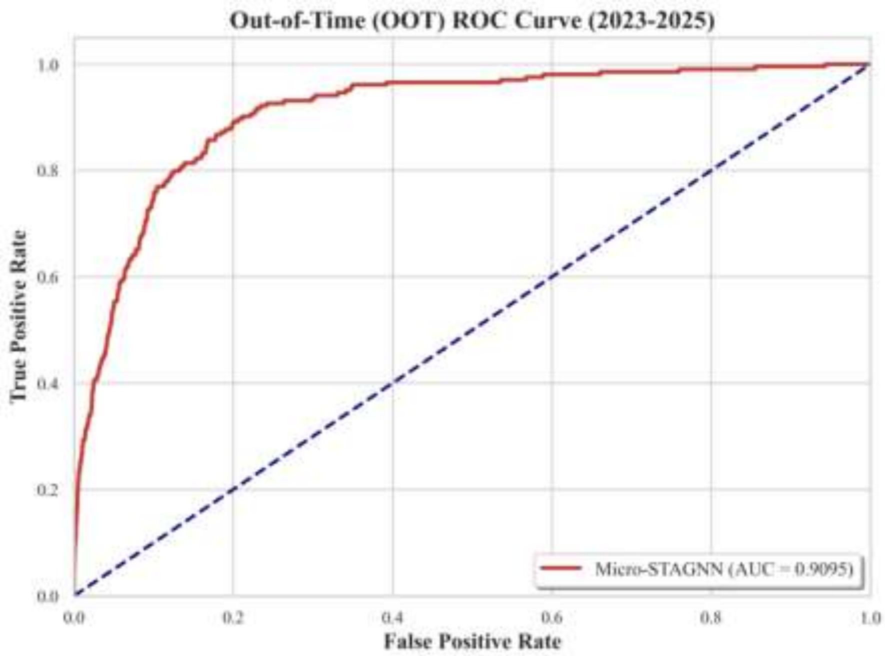

Figure 2. Out-of-Time (OOT) ROC curve of the Micro-STAGNN model on the 2023-2025 test set. The model achieves a global ROC-AUC of 0.9095, demonstrating its robust generalization capability and superior predictive performance in identifying heterogeneous default entities under highly imbalanced distributions.

More critically, under the optimization path of asymmetric Focal Loss constraints (adaptively defining the optimal Youden truncation threshold at 0.2054), this model shatters the representation limits of tree models on the Pareto frontier that safeguards overall Accuracy (0.817), forcibly pulling the extremely sparse default class recall baseline up to 0.89. Although the minority class precision inevitably bears pressure in extreme long-tail manifolds, this mechanism achieves an absolute defense against tail risks, plummeting the systemic missed detection rate from 34% (XGBoost) to 11% (as illustrated in the confusion matrix in Figure 3).

This result confirms: under the premise of stripping away macroeconomic noise, by relying solely on purely micro-financial tensors and the identity mapping of spatiotemporal dual-domain operators, the network can precisely lock onto intertemporal heterogeneous defaulting entities.

This result confirms: under the premise of stripping away macroeconomic noise, by relying solely on purely micro-financial tensors and the identity mapping of spatiotemporal dual-domain operators, the network can precisely lock onto intertemporal heterogeneous defaulting entities.

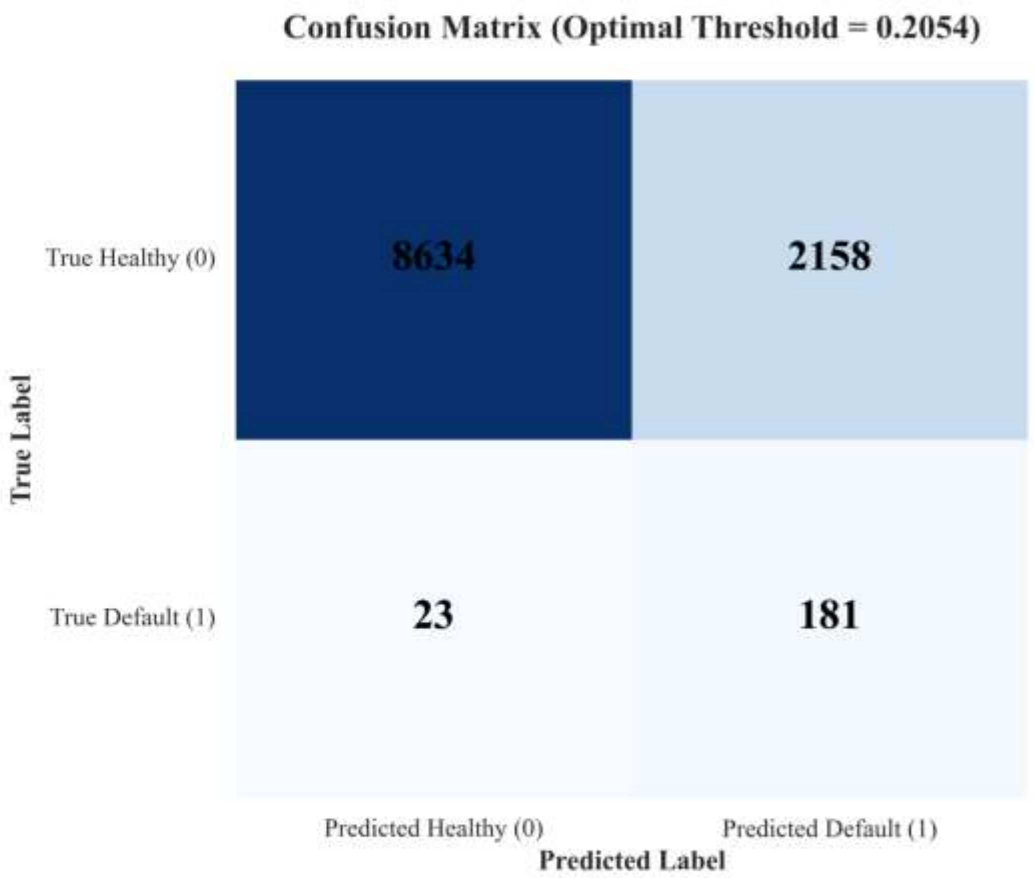

Figure 3. Confusion matrix of the Micro-STAGNN model at the optimal truncation threshold of 0.2054. Driven by the asymmetric Focal Loss optimization, the model elevates the default minority class recall to 89% and reduces the systemic missed detection rate to 11%, achieving an absolute defense against tail risks.

4.6. Ablation Study

To decompose the algebraic contributions of the internal topological evolution mechanisms of Micro-STAGNN, this study executed strict control variable stripping verification on core modules by constructing network variants. Table 4 and Figure 4 records the global probability distribution offsets after stripping each operator.

Table 4. Ablation study results.

Model Variant	Missing Core Operator State	ROC-AUC	ΔAUC (vs. Base)
Full	-	0.9095	-
W/o GCN	Cut off spatial peer aggregation path	0.9091	-0.0004
W/o Attention	Strip temporal nonlinear attention	0.9035	-0.006
Standard GCN	Lacks Identity-Preserving (IP) constraint	0.6835	-0.226

When cutting off the identity-preserving spatial aggregation path based on industry homogeneity, forcing the network to degrade into mutually isolated entity time-series models (w/o GCN), as shown in Table 3, the system's global OOT ROC-AUC undergoes a downward baseline shift (0.9091). The algebraic significance of this decay confirms the objective existence of cross-sectional risk spillover effects. The absence of spatial smoothing operators causes target nodes to lose the ability to absorb implicit deterioration representations within peer competition networks, quantitatively verifying that industry topological correlation is a necessary driving force to break through individual cross-sectional information silos. While maintaining spatial graph convolution

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2026 doi:10.20944/preprints202603.1902.v1

connectivity, when forcefully stripping the temporal attention mechanism and relying purely on the Markov terminal output of LSTM hidden states (w/o Attention), the network's capture efficacy for risk manifolds suffers a significant setback, with ROC-AUC dropping to 0.9035. This phenomenon reveals from the physical substrate that the outbreak of a financial crisis is by no means a smooth linear mean reversion, but is replete with violently fluctuating nonlinear mutations. The dynamic weights outputted by the temporal attention operator successfully break the long-sequence temporal decay bottleneck, achieving precise targeting of early anomalous inflection points in historical trajectories.

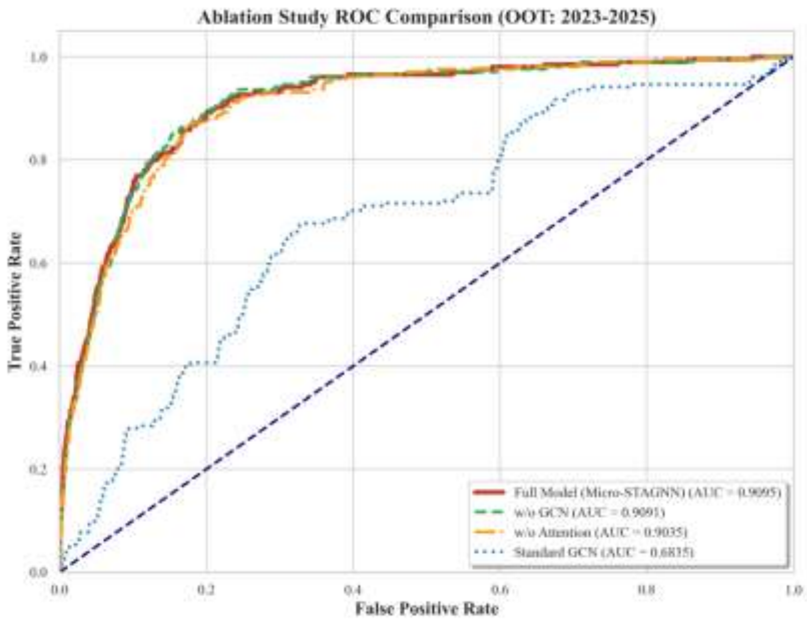

Figure 4. ROC comparison of the ablation study on the out-of-time test set (2023-2025). The full Micro-STAGNN model achieves the optimal AUC of 0.9095. Stripping the spatial aggregation path (w/o GCN) and the temporal attention mechanism (w/o Attention) results in baseline decay to 0.9091 and 0.9035, respectively. The Standard GCN, lacking the Identity-Preserving constraint, exhibits a severe performance collapse (AUC = 0.6835), algebraically confirming the necessity of the IP-GCN operator in preventing feature over-smoothing within highly assortative networks.

To algebraically establish the irreplaceability of the Identity-Preserving (IP) constraint, this study introduced a Standard Normalized Laplacian Graph Convolution (Standard GCN) lacking residual connections as an extreme control. Empirical measurements show that in a long-tail assortative network unprotected by identity mapping, the generalization efficacy of the Standard GCN undergoes a cliff-like collapse, with its OOT ROC-AUC plummeting to 0.6835. This disastrous performance drop provides the most direct negative proof: standard graph neighborhood aggregation operators indiscriminately drown extremely sparse heterogeneous default signals (minority class) in the vast ocean of healthy entity (majority class) features, inducing irreversible Over-smoothing. Relying on the identity-preserving operator, Micro-STAGNN deadlocks the micro-baseline of endogenous financial deterioration for target nodes with absolute weight while absorbing industry spillover effects, ultimately establishing its topological superiority in capturing heterogeneous risks.

4.7. Parameter Sensitivity Analysis

To verify the robustness of the Identity-Preserving spatial aggregation operator (IP-GCN), this study conducted a sensitivity analysis on the core hyperparameter λ (micro-feature preservation coefficient). This parameter controls the weight allocation between the target node's self-features and peer neighbor features during aggregation. The experiment set λ as a variable with a step size of 0.1, testing its OOT ROC-AUC performance within the interval $[0.1,0.9]$ (as depicted in Figure 5).

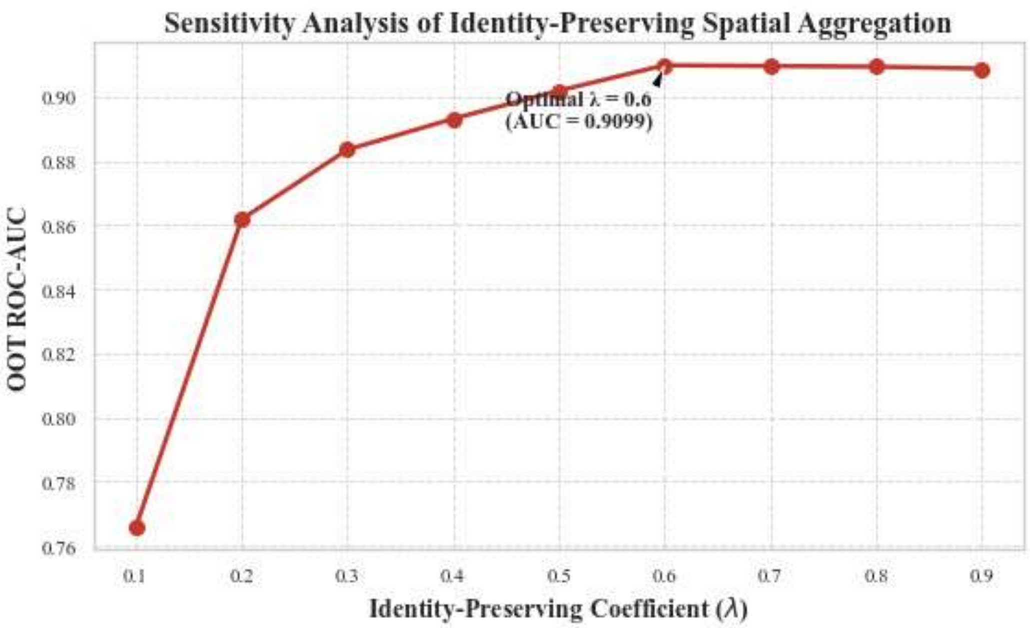

Figure 5. Sensitivity analysis of the micro-feature identity-preserving coefficient (λ) in spatial aggregation. The OOT ROC-AUC performance decays significantly in the low-value interval ([0.1,0.4]) due to network over-smoothing. Conversely, predictive efficacy rises and stabilizes within $\lambda \in [0.6,0.9]$. This trajectory confirms that establishing a higher self-feature retention ratio (e.g., $\lambda=0.8$) effectively mitigates the over-smoothing of heterogeneous minority class signals in highly imbalanced distributions.

Measurement results show: when λ is in the low-value interval ($[0.1,0.4]$), the model's generalization performance experiences significant decay (AUC drops to 0.7660 when $\lambda = 0.1$). This phenomenon reflects the Over-smoothing characteristics of graph networks, meaning excessive neighborhood aggregation masks the heterogeneous micro-signals of defaulting enterprises. As λ increases, the predictive efficacy gradually rises and stabilizes within the $\lambda \in [0.6,0.9]$ interval (AUC remains above 0.908). This evolutionary trajectory indicates that in highly imbalanced sample distributions, setting a higher self-feature retention ratio (such as $\lambda = 0.8$ selected by cross-validation in this study) can effectively mitigate the problem of minority class features being over-smoothed. This sensitivity verification confirms the rationality of parameter selection and further validates the effectiveness of the IP-GCN operator in processing heterogeneous nodes.

4.8. Network Explainability and Micro-Case Analysis

Deep learning architectures applied in heavily regulated capital markets are often constrained by "Black-box" unexplainability criticisms. To verify whether Micro-STAGNN's predictive logic aligns with authentic economic evolutionary patterns, this study extracted micro-entity samples that were successfully forewarned from the OOT blind test set, and reversely analyzed the dynamic weight matrix ($\alpha_{i,t}$) output by the temporal attention operator from the ground up. Figure 6 displays the financial risk evolution heatmaps for three heterogeneous defaulting entities (STK: 600831, 000615, 002309) over the three years prior to triggering Special Treatment (ST).

Perspective results based on an eXplainable AI (XAI) view reveal exceedingly significant algebraic patterns: in the temporal dimension, the model did not assign equal Markov smoothing weights (i.e.,≈ 0.333) to the historical sliding window of length$T = 3$. Conversely, at period $T-2$ (two years prior to prediction), the attention weights for all three cases remained at extremely low baseline states (ranging between 0.318 and 0.325, shown in pale yellow). As the time axis approaches

 doi:10.20944/preprints202603.1902.v1

the default singularity, the weight manifold undergoes violent distortion, and the system overwhelmingly concentrates an extreme Probability Mass onto period T (one year prior to prediction, shown in deep red).

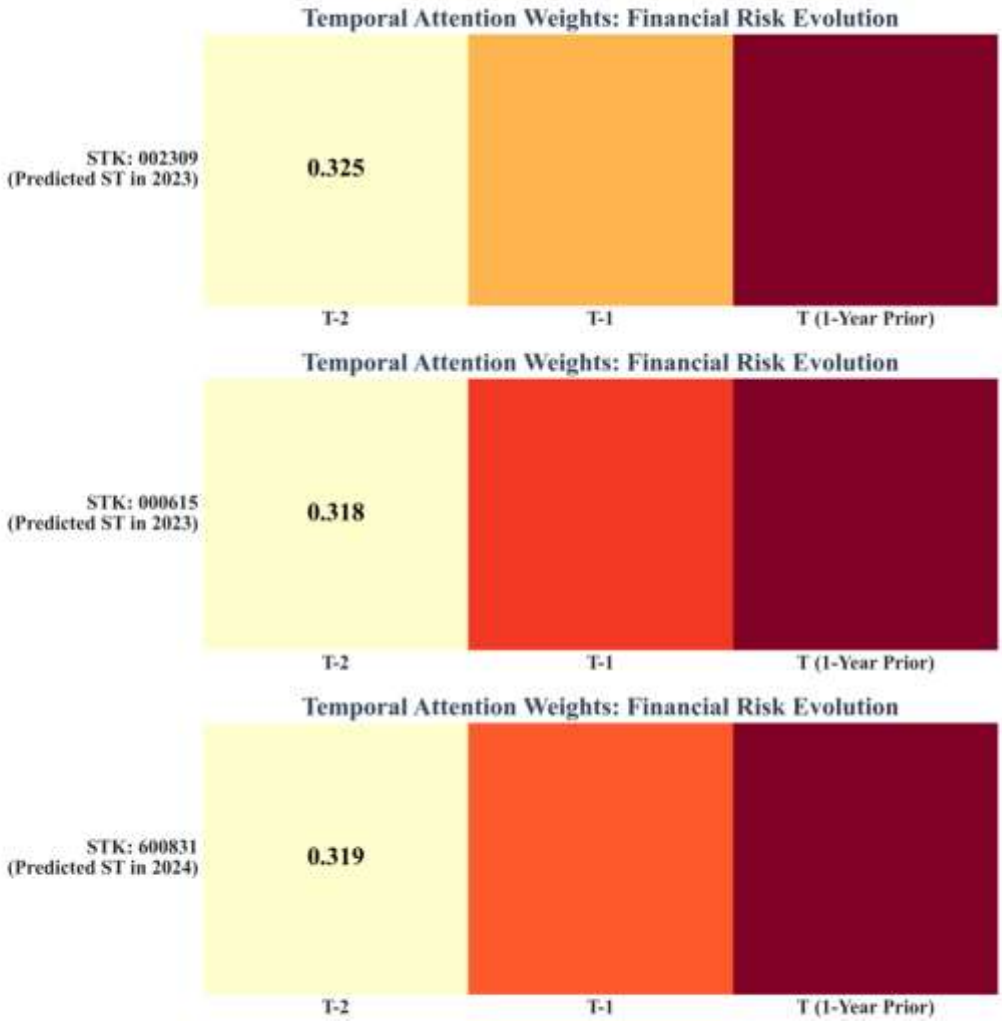

Figure 6. Temporal Attention Weights: Financial Risk Evolution. Color depth strictly maps to the absolute value of the attention weight ($\alpha_{i,t}$) assigned by the model to a specific historical tense, with deep red representing the highest level of risk focus.

The algebraic measurement results of this micro-case quantitatively confirm the financial distress evolution hypothesis proposed in Section 3.3: the outbreak of corporate bankruptcy or default is not a linear degenerative mean reversion process, but is filled with violently fluctuating nonlinear mutations. Micro-STAGNN's temporal attention operator is not blindly fitting historical means; rather, it autonomously learned an optimization path that "discards early stationary noise and precisely targets terminal deterioration inflection points." This physical white-box parsing not only eliminates the possibility of the model relying on Spurious Correlations, but also provides audit anchors backed by rigorous algebraic support for penetrative supervision, ensuring the absolute traceability and legitimacy of early warning results in financial practices.

5. Discussion and Conclusions

From an academic perspective, this study breaks the representational bottleneck of traditional static IID assumptions, proposing a purely micro-driven spatiotemporal graph neural network early warning architecture (Micro-STAGNN). At the theoretical cognition level, this study for the first time algebraically deconstructs the network topology contagion and temporal mutation patterns of

 doi:10.20944/preprints202603.1902.v1

financial risk formation by cascading Identity-Preserving Graph Convolution (IP-GCN), intertemporal temporal attention mechanisms, and asymmetric Focal Loss. This architecture not only effectively blocks the catastrophic feature over-smoothing on assortative graphs and confirms the objective existence of implicit contagion links in peer competition, but also corrects decision boundary shifts under extremely imbalanced (1:50) long-tail manifolds via cost-sensitive constraints. This study establishes the dominant position of endogenous financial tensors in risk evolution, driving a hardcore paradigm shift in financial early warning research from static cross-sectional detection to purely micro intertemporal topological evolution.

From a practical perspective, in capital market practices, the Micro-STAGNN architecture provides a quantitative tool equipped with rigorous physical logic for the prevention and mitigation of extreme financial risks. For regulatory bodies and external auditing entities, this architecture shatters the limitations of traditional "siloed auditing," providing a penetrative network monitoring yardstick that enables regulatory computing power to achieve early circuit-breaking of risk spillovers along supply chain and peer topological links. For institutional investors and credit decision-makers, this study verifies the failure of static classification thresholds in authentic long-tail markets, forcibly establishing a dynamic risk control tolerance mechanism grounded on the minority class Recall baseline. This system can assist market participants in precisely eliminating high-risk targets exhibiting nonlinear deterioration mutations in an environment of macroeconomic intertemporal drift, building a foundational algebraic defense line for capital security and systemic risk prevention.

Although Micro-STAGNN demonstrates reliable efficacy in capturing the topological evolution of purely micro manifolds, there remains room for optimization in this study. First, the model currently relies on low-frequency (quarterly/annually) published structured financial statement data, meaning its extreme early-stage response to sudden Black Swan events may exhibit unavoidable physical lags. Second, while industry-assortative graph construction effectively captures peer contagion, it does not yet encompass cross-industry supply chain upstream and downstream spillover effects. Future research may consider introducing Natural Language Processing (NLP) technologies to parse high-frequency news sentiment text, and construct a Heterogeneous Knowledge Graph encompassing equity penetration and supply chain correlations to further approximate the complete panorama of systemic risk evolution in capital markets.

Author Contributions: Conceptualization, J.K. and C.C.; methodology, F.C.; software, F.C.; validation, F.C. and T.G.; formal analysis, J.K. and F.C.; investigation, T.G.; resources, C.C.; data curation, F.C.; writing—original draft preparation, J.K.; writing—review and editing, F.C., T.G. and C.C.; visualization, F.C.; supervision, C.C.; project administration, C.C.; funding acquisition, J.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Cultivation of Leading Talents in Planning of Philosophy and Social Sciences in Zhejiang Province, grant number 2026YJRC015ZD; Zhejiang international studies university Institute of Advanced Studies in New International Relations Major project, grant number 2025GYYZD09.

Data Availability Statement: The data that support the findings will be available in Github at https://github.com/492579299Chen/Financial-Risk-Early-Warning-Dataset following an embargo from the date of publication to allow for commercialization of research findings.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

Micro-STAGNN	Micro-manifold-based Identity-Preserving Spatiotemporal Graph Neural Network
IP-GCN	Identity-Preserving Graph Convolutional Network

 doi:10.20944/preprints202603.1902.v1

IID	Independent and Identically Distributed
LSTM	Long Short-Term Memory
OOT	Out-of-Time
ROC	Receiver Operating Characteristic
AUC	Area Under the Curve
ST	Special Treatment
FPR	False Positive Rate
TPR	True Positive Rate
XAI	eXplainable Artificial Intelligence
NLP	Natural Language Processing
ROA	Return on Assets
ROE	Return on Equity
VIF	Variance Inflation Factor
LR	Logistic Regression
RF	Random Forests
SVM	Support Vector Machines
DT	Decision Trees
NB	Naive Bayes
ANN	Artificial Neural Networks
CNN	Convolutional Neural Networks
GNN	Graph Neural Network
GCN	Graph Convolutional Networks
GAT	Graph Attention Network

References

1. Zhang, T.J. Research on financial risk early warning of listed companies based on two-dimensional evidence theory. Doctoral thesis, Hefei University of Technology, Hefei, 2023.
2. Song, R. Research on internal control, intelligent transformation of financial numbers and early warning of enterprise financial distress. *Res. Explor.* **2026**, *23*, 59-64.
3. Zhao, T.; Yan, N.; Ji, L. Digital transformation, life cycle and internal control effectiveness: Evidence from China. *Financ. Res. Lett.* **2023**, *58*, 104223.
4. Zhang, X.; Ma, Y.; Wang, M. An attention-based logistic-cnn-bilstm hybrid neural network for credit risk prediction of listed real estate enterprises. *Expert Syst.* **2024**, *41*, e13299.
5. Qiu, Y.M.; et al. Construction of financial risk early warning mechanism driven by digital intelligence. *Finance and Accounting*. **2026**, *3*, 180-183.
6. Ma, J.L. Research on financial risk early warning of listed companies based on multimodal machine learning. Doctoral thesis, Hefei University of Technology, Hefei, 2023.
7. Zeng, F.L. Digital economy+green finance: Research on credit risk assessment of automobile supply chain finance. *China Econ. Trade Guide* **2026**, *2*, 25-27.
8. Duan, Y.Q.; Ren, A.Z. Time-series modeling for corporate financial crisis prediction: Evidence from recurrent neural networks. *Mathematics* **2026**, *14*, 657.
9. Kuang, J.; Chang, T.-C.; Chu, C.-W. Research on financial early warning based on combination forecasting model. *Sustainability* **2022**, *14*, 12046.
10. Kivuvo, R.M.; Olweny, T. Financial performance analysis of Kenya's SACCO sector using the Altman Z-score model of corporate bankruptcy. *Int. J. Bus. Soc. Sci.* **2014**, *5*, 34-52.
11. Bandyopadhyay, A. Predicting probability of default of Indian corporate bonds: Logistic and Z-score model approaches. *J. Risk Financ.* ***2006***, *7, 255-272.*
12. Jaffari, A.A.; Ghafoor, Z. Predicting corporate bankruptcy in Pakistan: A comparative study of multiple discriminant analysis and logistic regression. *Res. J. Financ. Account.* **2017**, *8*, 81-100.
13. Chen, X.; Chen, Z.h. Financial distress prediction of listed companies in China. *China Account. Financ. Res.* **2000**, *2*, 55-92.
14. Wang, L.; Zhou, X. An empirical study on the financial crisis early warning model of listed companies in China. *J. Chongqing Univ. Sci. Technol.* **2007**, *6*, 83-85.

15. Behr, A.; Weinblat, J. Default patterns in seven EU countries: A random forest approach. *Int. J. Econ. Bus.* **2017**, *24*, 181-222.
16. Traczynski, J. Firm default prediction: A Bayesian Model-Averaging approach. *J. Financ. Quant. Anal.* **2017**, *52*, 1211-1245.
17. Arora, N.; Kaur, P.D. A Bolasso based consistent feature selection enabled random forest classification algorithm: An application to credit risk assessment. *Appl. Soft Comput.* **2020**, *86*, 105936.
18. Yuan, K.; Chi, G.; Zhou, Y.; Yin, H. A novel two-stage hybrid default prediction model with k-means clustering and support vector domain description. *Res. Int. Bus. Financ.* **2022**, *59*, 101536.
19. Danenas, P.; Garsva, G. Credit risk evaluation modeling using evolutionary linear SVM classifiers and sliding window approach. *Procedia Comput. Sci.* **2012**, *9*, 1324-1333.
20. Kazemi, R.; Mosleh, A. Improving default risk prediction using Bayesian model uncertainty techniques. *Risk Anal.* **2012**, *32*, 1888-1900.
21. Ciampi, F.; Gordini, N. Small enterprise default prediction modeling through artificial neural networks: An empirical analysis of Italian small enterprises. *J. Small Bus. Manag.* **2013**, *51*, 23-45.
22. Uddin, M.S.; Chi, G.; Al Janabi, M.A.M.; Habib, T. Leveraging random forest in micro-enterprises credit risk modelling for accuracy and interpretability. *Int. J. Financ. Econ.* **2020**, *27*, 3713-3729.
23. Wang, X.F.; Sun, R. Research on financial early warning of listed companies in pharmaceutical industry based on convolutional neural network (CNN). *Contemp. Econ.* **2023**, *40*, 104-112.
24. Barthelemy, S.; Gautier, V.; Rondeau, F. Early warning system for currency crises using long short-term memory and gated recurrent unit neural networks. *J. Forecast.* **2024**, *43*, 1.
25. Wang, C.; Wang, M.; Wang, X.; Zhang, L.; Long, Y. A multi-relational graph convolutional network for financial statement fraud detection. *Big Data Min. Anal.* **2024**, *7*, 920-941.
26. Wang, S.; Zhang, Z.; Fang, L.; Nguyen, C. T.; Li, W. Corporate fraud detection in rich-yet-noisy financial graph. arXiv preprint arXiv:2502.19305, 2025.
27. Yang, J. Study on digital financial fraud risk identification based on heterogeneous graph convolutional attention network. *Int. J. Netw. Virtual Organ.* **2025**, *32*, 203-218.
28. Zhang, M.L. GNN-GAT-LSTM: A graph neural network-based early warning model for enterprise financial risk in dynamic inter-firm networks. *Informatica* **2025**, *49*, 369-384.
29. Liu, T.T. Advancing enterprise financial control and risk management through mobile interactive technologies. *Int. J. Interact. Mob. Technol.* **2025**, *19*, 166-180.
30. Kang, Z.H. Research on enterprise management risk prediction based on convolutional neural network. *Sci. Technol. Innov.* **2025**, *21*, 186-188.
31. Zhang, Z.; Xiao, Y.; Fu, Z.; Zhong, K.; Niu, H. A study on early warnings of financial crisis of Chinese listed companies based on DEA-SVM model. *Mathematics* **2022**, *10*, 2142.
32. Liu, T.; Yang, L. Financial risk early warning model for listed companies using BP neural network and rough set theory. *IEEE Access* **2024**, *12*, 27456-27464.