Transactions on Computational and Scientific Methods | Vo. 4, No. 4, 2024
ISSN: 2998-8780
https://pspress.org/index.php/tcsm
Pinnacle Science Press

Graph Neural Network Framework for Default Risk Identification in Enterprise Credit Relationship Networks

Yuxiu Lin
Columbia University, New York, USA
yl4865@columbia.edu

Abstract: This paper proposes a graph neural network-based algorithm for default risk identification in enterprise credit relationship networks, addressing the challenges of complex dependency structures and multi-source feature integration. The method constructs a graph representation based on enterprise guarantee relationships, integrating node-level information such as financial attributes, registration details, and credit labels. A multi-layer graph convolutional network combined with a multi-head attention mechanism is employed to capture structural dependencies and aggregate features. The model architecture comprises a graph embedding and a classification module for performing binary classification between default and non-default enterprises. To evaluate performance, multiple experiments are conducted, including structure perturbation, training ratio variation, layer depth, and attention head configuration. The proposed model is assessed against mainstream methods using metrics such as accuracy, F1-score, and precision. Results show that the model outperforms existing baselines in structural representation, risk identification accuracy, and robustness to graph noise. Especially in dense credit graphs with multi-hop dependencies, the model effectively captures deep relational patterns and abnormal paths, demonstrating strong classification performance and stability for credit risk tasks in complex financial networks.

Keywords: graph neural network; enterprise credit graph; default identification; structural modeling

1. Introduction

In modern financial systems, corporate credit risk management remains a core issue[1]. As financing structures become increasingly complex and credit entities diversify, traditional credit assessment methods based on financial indicators and static scoring systems face significant challenges. Enterprises build complex credit relationship networks through trade, investment, guarantees, and other forms of interaction. These network structures influence not only the creditworthiness of individual firms but also pose systemic risk through potential contagion effects. Therefore, identifying potential default-prone entities within large-scale heterogeneous enterprise data has become a critical technical challenge in financial regulation and risk control[2].

Compared with traditional methods, graph neural networks (GNNs) have inherent advantages in modeling non-Euclidean structured data. They can capture the graph-based characteristics of credit relationships among enterprises. This approach leverages not only the features of individual enterprise nodes but also indirect information from connected entities, thus offering a more comprehensive semantic view of the credit network. For instance, a change in the credit status of a supplier or customer may signal risk that propagates through the graph. GNNs are capable of learning from such relational paths. On this basis, discriminative algorithms built on credit graphs can go beyond the limits of static models and improve the modeling of credit risk dynamics.

The advancement of financial technology provides new opportunities for structured relationship modeling. Techniques such as graph representation learning, graph attention mechanisms, and multi-hop path extraction have made it feasible to model complex credit structures among enterprises efficiently. Corporate default is rarely an isolated event. A combination of network structure, industry linkages, and historical cooperation often influences it. A GNN-based modeling framework enables deep abstraction and generalization of credit dependencies, supporting more accurate risk identification and classification. This is of practical value in credit issuance, investment management, and regulatory decision-making[3].

Moreover, credit risk in enterprises is inherently a temporal and group-coupled classification problem. Traditional node classification or regression models struggle to capture their evolving nature. GNNs, through structure-aware mechanisms and embedded representations, offer an efficient way to compress and integrate high-dimensional credit behavior. This helps reduce overfitting caused by high feature dimensionality and allows for modeling long-term credit trends using historical behavior data. Risk identification models based on such frameworks can retain micro-level features while capturing global associations in the credit network. This leads to better robustness in identifying financial entities with vague risk boundaries and complex relationships[4].

In summary, corporate credit risk management is shifting from static analysis toward structure-aware and dynamic modeling. Against this background, research on GNN-based credit relationship modeling and risk classification represents not only a technical frontier but also a vital step toward intelligent and systematic financial risk control. By deeply analyzing enterprise credit networks and extracting their graph structures, it becomes possible to enhance the accuracy, sensitivity, and interpretability of risk identification. This lays a solid theoretical and methodological foundation for building a more transparent, fair, and efficient credit evaluation system.

2. Relevant Literature

Corporate credit risk identification has long been a key research topic in the financial domain. Traditional approaches rely on financial statements, credit scores, and historical default records. Classical machine learning algorithms such as logistic regression, decision trees, and support vector machines are commonly used. These methods perform reasonably well when data is structured and feature relationships are clear[5]. However, they often struggle with generalization and interpretability when facing complex, nonlinear credit behaviors in real-world settings. Moreover, static modeling overlooks interactions between enterprises and ignores the dynamic external environment. This limits the ability to capture risk transmission mechanisms and systemic vulnerabilities[6].

In recent years, as modeling capabilities for graph-structured data have improved, researchers have turned their attention to credit network relationships among enterprises. They attempt to extract risk features from structural information. Some studies apply graph theory tools by constructing guarantee graphs or transaction graphs. They use graph metrics such as degree centrality, clustering coefficient, and shortest path length for risk identification. While these methods broaden the scope of risk representation, they still rely on handcrafted features. This makes it difficult to automatically learn deep representations from graph structures. With the rise of graph neural networks, a more expressive and automated modeling paradigm has been introduced into credit risk prediction tasks. This has become a powerful alternative to traditional graph-based analysis[7].

Graph neural networks use multi-layer message passing mechanisms to integrate information from nodes and their neighbors. They can effectively model multi-hop dependencies and contextual information in graph structures. Existing studies have applied GNNs to scenarios such as banking guarantee networks and supply chain networks. They use graph convolution or attention mechanisms to extract representation vectors of enterprises within their credit networks for default prediction and credit scoring. These models incorporate not only the financial and behavioral features of each node but also the global structure of the credit graph. This improves their ability to recognize complex credit behavior. Some studies have further introduced heterogeneous graph modeling strategies. These integrate different types of enterprise relationships and multi-source data, offering new perspectives for multidimensional credit risk assessment[8].

Meanwhile, the combination of GNNs with sequence modeling has emerged as a new direction. Some research explores integrating time series features with graph structures to handle the evolving nature of corporate credit. In such models, GNNs capture static or slowly changing structural dependencies, while recurrent networks or transformer architectures are used to model behavioral trajectories over time. This fusion significantly enhances sensitivity to sudden credit risk changes and chain reactions. It is especially useful in multi-period, cross-industry, and highly interconnected financial environments. Overall, existing work has provided a rich foundation for credit relationship modeling and risk identification. However, limitations remain in terms of model generalization, coarse graph definitions, and insufficient integration of heterogeneous information. Further improvements are needed to enhance both expressiveness and practical utility.

3. Method Overview

The enterprise credit relationship modeling and default risk identification algorithm proposed in this paper is based on the graph neural network framework. The core idea is to build an enterprise credit relationship graph and combine node characteristics and structural dependencies to perform end-to-end risk identification modeling. Its model architecture is shown in Figure 1.

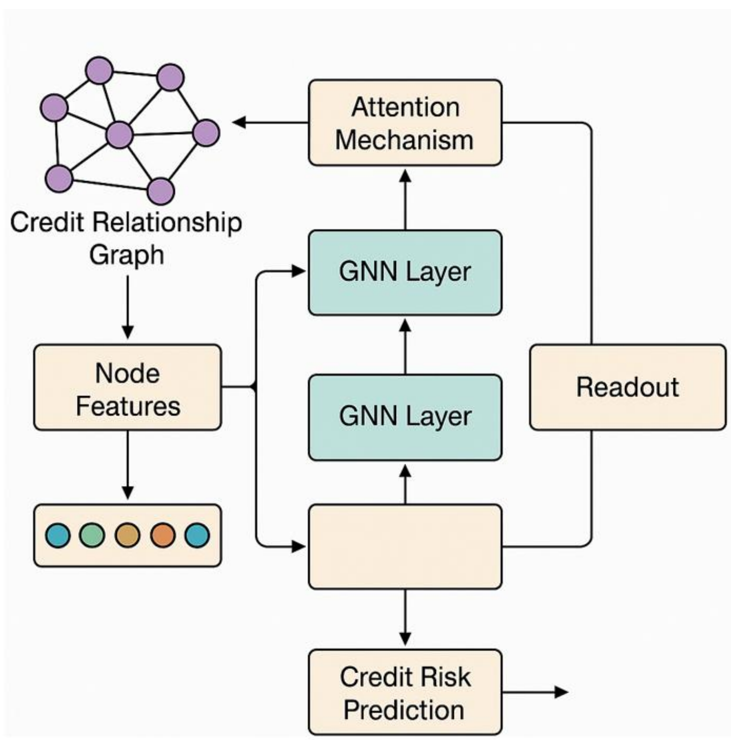

Figure 1. Architecture of the GNN-based Default Risk Identification Framework

First, let the enterprise set be $V = \{v_1, v_2, ..., v_N\}$, each node represents an enterprise, and the graph structure is represented by $G = (V, \varepsilon)$, where $\varepsilon \subseteq V \times V$ represents the credit relationship between enterprises, such as guarantee, transaction, investment, etc. Each node v_i corresponds to a feature vector $x_i \in R^d$, which represents its multi-dimensional information, such as financial, behavioral, and industry attributes.

In the representation learning process of graph neural networks, the model updates the representation of the central node by aggregating the feature information of neighboring nodes. The update formula of the basic graph convolution layer can be expressed as:

$$h_i^{(l+1)} = \sigma(\sum_{j \in N(i)} \frac{1}{\sqrt{|N(i)||N(j)|}} W^{(l)} h_j^{(l)})$$

$N(i)$ represents the neighbor set of node i, $h_j^{(l)}$ represents the representation of node j in the lth layer, $W^{(l)}$ represents the learnable weight, and σ represents the nonlinear activation function. This structure can encode the local graph topology and node features, and realize the structural perception of corporate credit semantics.

In order to further improve the model's ability to distinguish heterogeneous relationships between enterprises, the attention mechanism is introduced to perform weighted aggregation on neighboring nodes. The attention weights of node i and neighbor j are calculated as follows:

$$\alpha_{ij} = \frac{\exp(LeakyRELU(a^T[Wh_i \| Wh_j]))}{\sum_{k \in N(i)} \exp(LeakyRELU(a^T[Wh_i \| Wh_k]))}$$

Where a is the attention parameter vector, || represents the vector concatenation operation, and W is the feature transformation matrix. This mechanism can automatically learn the strength of relationships and highlight the contribution of important connections to risk transmission.

After obtaining the final node embedding vector z_i, the default probability mapping model is performed through the fully connected layer, and the discriminant function form is as follows:

$$\hat{y}_i = \sigma(w^T z_i + b)$$

Where w and b are classifier parameters, and $\hat{y}_i \in (0,1)$ represents the probability of node i defaulting. To optimize the overall model, a weighted cross-entropy loss function is used for training, which is:

$$L = -\sum_{i=1}^{N} w_i [y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i)]$$

$y_i \in \{0,1\}$ is the true default label, and w_i is the sample weight, which is used to balance the imbalance of positive and negative samples.

The overall algorithm architecture supports a unified optimization process of multi-layer graph embedding learning, relationship-weighted perception, and risk output mapping. This method not only can model static information of enterprises, but also can strengthen the modeling expression of credit risk transmission paths through graph structures, providing structured support for the identification of high-risk enterprises. The model adopts an end-to-end back propagation method at the full graph level during training to achieve joint optimization of feature learning, structural modeling, and risk discrimination.

4. Experimental Dataset

This study uses the Tianchi Enterprise Guarantee Network Dataset. The dataset is provided by the Alibaba Tianchi platform and is widely used in corporate credit risk analysis and graph-based modeling tasks. It contains a large number of enterprise nodes and their mutual guarantee relationships, forming a graph structure with clear economic significance. It is well-suited for studying how credit relationships influence default risk.

The dataset is organized in the form of a graph. Each node represents an enterprise. Each edge indicates a guarantee relationship between two enterprises. Edges may have direction and weight attributes. In addition, each node is associated with multi-dimensional features such as registered capital, years of operation, industry type, and credit rating. These features support the construction of joint modeling frameworks based on both node attributes and structural information. The dataset also distinguishes between default and non-default enterprises, making it suitable for supervised learning tasks.

To ensure experimental feasibility and real-world relevance, the raw data were cleaned, deduplicated, and structured. A stable graph and corresponding feature matrix were constructed. The final graph contains about 50,000 enterprise nodes and nearly 100,000 guaranteed edges. It forms a high-density credit network with multiple interwoven paths. This structure is appropriate for deep modeling and risk classification using graph neural networks.

5. Results and Analysis

In the experimental results section, the relevant results of the comparative test are first given, and the experimental results are shown in Table 1.

Table 1: Comparative experimental results

Method	Accuracy	Macro-F1	Precision
GDAN[9]	84.5%	83.1%	81.6%
TemGNN[10]	83.2%	81.7%	80.2%
CCR-GNN[11]	81.8%	80.4%	78.9%
SparseGraphSage[12]	80.6%	79.1%	77.3%
Ours	86.3%	84.9%	83.8%

Overall, the proposed model demonstrates superior performance in corporate default risk classification. In terms of Accuracy, the model achieves 86.3%, nearly two percentage points higher than the best-performing baseline model, GDAN. This shows that incorporating graph structure awareness and relationship modeling significantly improves overall classification accuracy. The result indicates that the model can more accurately distinguish between default and non-default enterprises, reflecting stronger risk identification capability.

For the Macro-F1 score, the model also achieves the highest value of 84.9%, outperforming all comparison models. This suggests that it maintains strong classification ability even under class imbalance. Compared with models such as CCR-GNN and TemGNN, the improvement in F1 score means the model not only better identifies default enterprises (the minority class) but also reduces false positives for non-default firms. This reflects the robustness and adaptability of the model in credit risk classification tasks.

In terms of Precision, the model reaches 83.8%, significantly outperforming SparseGraphSage at 77.3% and CCR-GNN at 78.9%. A higher Precision indicates that a greater proportion of enterprises predicted as default are indeed defaulting. This reduces the risk of false alarms. In real financial risk control scenarios, this capability is important, as misclassifying non-default firms as high-risk can directly impact credit issuance and regulatory decisions.

Taken together, the model achieves strong performance across all three metrics. The proposed graph neural network-based credit relationship modeling algorithm not only excels in accuracy but also shows balanced and reliable classification ability. Modeling enterprise guarantee and cooperation relationships as graph structures helps uncover multi-hop dependencies and hidden risk signals. This provides a more precise and systematic solution for intelligent default risk identification.

This paper also explores the impact of changing the number of graph convolution layers on the overall performance of the proposed model, focusing on how different layer depths affect the model's ability to learn and represent complex structural patterns within enterprise credit networks. In graph neural networks, the number of convolutional layers determines the receptive field of each node, influencing how far information can propagate through the graph and how deeply multi-hop dependencies are captured. Adjusting this architectural parameter plays a crucial role in balancing local feature extraction and global structural awareness, which are both essential for accurate risk classification. By systematically varying the number of layers, the study aims to analyze how network depth contributes to the expressiveness, stability, and generalization capacity of the model. The corresponding experimental setup and comparative findings are illustrated in Figure 2, providing a comprehensive view of this critical architectural factor in the context of graph-based credit risk modeling.

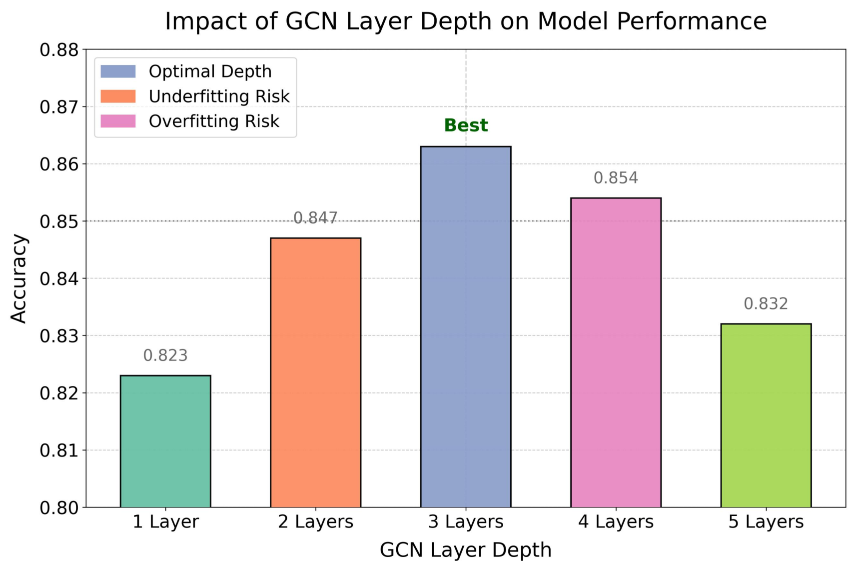

Figure 2. The impact of changing the number of graph convolution layers on model performance

The impact of the number of graph convolutional layers on model performance shows a clear trend. As the number of layers increases, the model's accuracy improves at first and reaches its highest value (86.3%) at three layers. After that, a slight decline occurs, showing a typical overfitting turning point. This trend suggests that increasing the network depth moderately enhances the model's ability to capture multi-hop dependencies and complex structures in the credit graph. It helps extract hidden risk relationships and improves classification accuracy.

When the model has only one or two layers, performance is weaker. This indicates that the graph information is not sufficiently propagated. Nodes can only perceive local features from immediate neighbors. As a result, the model cannot express complex credit relationships and shows clear underfitting risks. This outcome suggests that shallow structures are not adequate for capturing deep credit influence paths across levels and industries in the enterprise network.

When the graph convolutional depth exceeds three layers, such as with four or five layers, the model still performs well, but accuracy declines. This may be due to oversmoothing, where stacking too many layers leads to excessively similar node representations. The reduction in feature distinction weakens the model's classification boundaries and may cause overfitting or degraded expressiveness.

This paper also examines the impact of the number of attention heads on the representation ability of graph structure, focusing on how different configurations influence the model's capacity to capture and differentiate key relational patterns within the enterprise credit network. In graph neural networks equipped with attention mechanisms, the number of attention heads determines how many independent subspaces the model uses to attend to neighborhood information, affecting both the diversity and granularity of structural learning. By varying this parameter, the study investigates the balance between expressive power and noise sensitivity in structural representation. The corresponding experimental design and comparative analysis are illustrated in Figure 3, providing insight into the role of attention head settings in enhancing graph-based modeling of credit dependencies.

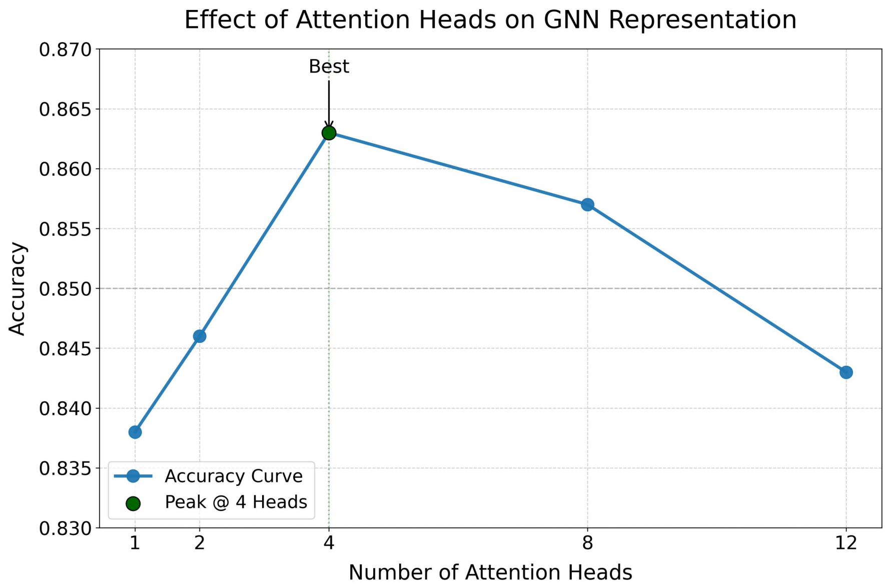

Figure 3. The impact of the number of attention heads on the representation ability of graph structures

The experimental results show that the number of attention heads has a significant impact on the performance of graph neural networks in modeling enterprise credit graph structures. When the number of attention heads is 4, the model achieves the highest accuracy (86.3%). This indicates that the model reaches an optimal balance in focusing on neighboring structures and capturing key credit dependencies through multi-channel attention. The result confirms the effectiveness of multi-head attention in enhancing structural representation and improving risk classification accuracy.

When the number of attention heads is low, such as 1 or 2, the model shows weaker performance. This suggests a limited focus range and insufficient representation capacity in structural modeling. In this case, semantic information from different neighbor paths is not well separated. As a result, credit dependencies among enterprises cannot be fully captured, which reduces the precision of risk identification.

When the number of attention heads increases further to 8 or 12, model performance starts to decline. This may be due to the accumulation of redundant information and overlap among subspaces. Such overlap can dilute feature representations and introduce structural noise, making it harder for the model to focus on truly important graph information. This “over-expression” effect is especially evident in complex and heterogeneous credit networks, reflecting the model's sensitivity to structural noise.

Therefore, setting the number of attention heads properly is essential for improving the resolution of graph structure information and enhancing the ability to model credit risk. The experimental results suggest that using four attention heads achieves the best structure-aware performance in enterprise credit network tasks. It maintains diversity in information channels while avoiding over-complication of structural expression, which helps improve the accuracy of identifying hidden credit risk.

This paper also gives an evaluation of the model's robustness under changes in the proportion of graph structure perturbations, and the experimental results are shown in Figure 4.

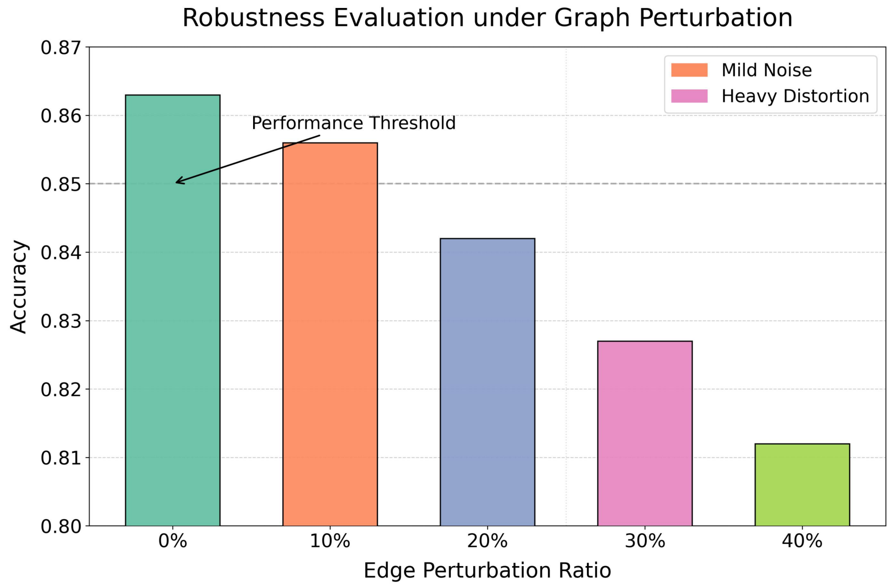

Figure 4. Model robustness evaluation under changing graph structure perturbation ratios

The experimental results show that the model performs best when the graph structure is undisturbed (0%), achieving an accuracy of 86.3%. This indicates that the model has strong risk identification capability on the original credit relationship network. It demonstrates that the proposed graph neural network can effectively model multi-hop dependencies and risk propagation paths between enterprises in an ideal, noise-free environment. The model provides an accurate structural representation of credit relationships.

When the disturbance ratio increases to 10%, the model remains relatively stable. Accuracy stays around 85.6%, with only a slight decrease. This suggests that the model has a certain level of robustness under mild structural perturbation. Even when some edges are altered or disrupted, the model can still capture key relationships in the graph and make reliable predictions on default risk. This reflects its ability to resist local structural noise.

However, as the disturbance ratio increases to 20%, 30%, and 40%, model performance declines significantly. Accuracy drops to 84.2%, 82.7%, and 81.2%, respectively. This trend indicates that the model relies heavily on the completeness of the graph structure. When a large proportion of edge information is disturbed, the model struggles to perceive the original credit paths and risk dependencies. This negatively impacts the prediction results. The findings highlight the critical role of credit graph structure in model classification performance.

In summary, the proposed model shows robust performance under mild graph disturbances but suffers noticeable degradation under severe structural interference. This experiment validates the boundary of the model's robustness in handling graph uncertainty and missing information in real-world settings. It also provides methodological insights for future improvements in abnormal structure repair and graph augmentation techniques.

This paper further investigates how variations in the size of the training set affect the model's generalization ability, which is a critical factor in the performance and stability of learning algorithms. In the context of enterprise credit risk modeling based on graph neural networks, the amount of training data directly influences the model's capacity to capture complex structural dependencies, learn discriminative representations, and adapt to unseen scenarios. A detailed exploration is conducted to examine how different proportions of training samples contribute to the learning process, particularly in graph-structured environments where node interactions and multi-hop relations play a central role. The corresponding experimental setup and observed trends related to training scale adjustments are illustrated in Figure 5, offering valuable insights into how training data availability shapes model robustness and predictive consistency.

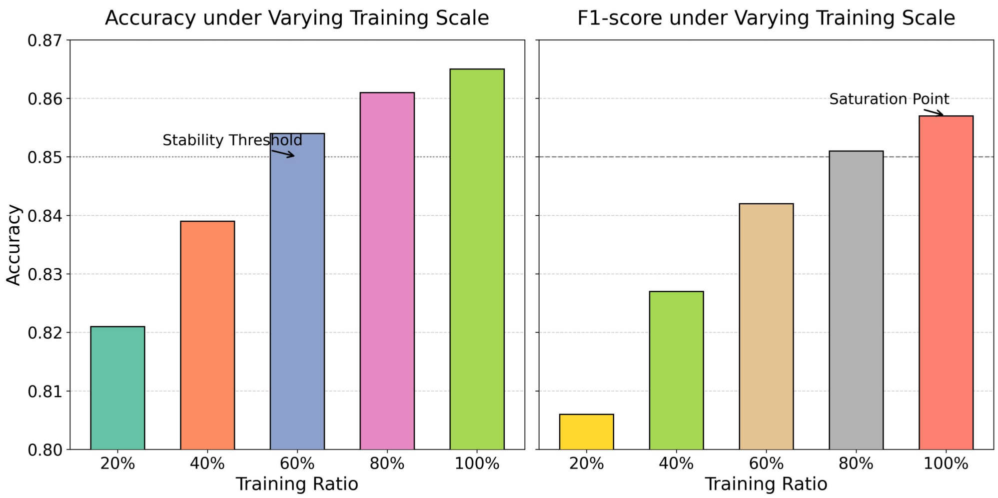

Figure 5. The impact of changes in training set size on model generalization ability

As shown in Figure 5, the size of the training set has a significant impact on the model's generalization ability. With the gradual increase in the training ratio, both Accuracy and F1-score show a consistent upward trend. This indicates that a larger training set helps the model better learn complex credit relationships and multi-hop dependencies among enterprises, thereby improving risk identification and classification accuracy.

In terms of Accuracy, when the training ratio increases from 20% to 60%, the model's performance improves significantly. Around 60%, the model reaches a "stability threshold," suggesting that once sufficient structural samples are provided, the model's ability to capture risk paths becomes stable. This phase marks a transition from underfitting to effective generalization and reflects enhanced understanding of the credit graph structure.

For the F1-score, the model's overall classification ability under class imbalance continues to improve as the training ratio increases. When the training ratio reaches 100%, the model hits a "saturation point," where further increasing the training size brings limited performance gain. This suggests that graph neural networks may reach a performance ceiling after absorbing structural knowledge. In dense graphs, overtraining can introduce redundancy rather than further improvements.

Overall, the experiment confirms that training set size is a sensitive factor in enterprise credit graph modeling. Proper control of the training ratio not only enhances performance but also balances training cost and practical deployment needs. In risk identification systems, the training scale should be selected based on data size and graph complexity to ensure stability and usability across different scenarios.

6. Conclusion

This paper presents a novel graph neural network-based approach for modeling enterprise credit relationships and identifying default risk. By integrating structural representation learning with node-level attribute modeling, the proposed framework effectively captures multi-hop dependencies and complex credit interactions within large-scale guarantee networks. The experimental results demonstrate that the model achieves superior classification accuracy and robustness compared to existing baseline methods, particularly under varying training scales, attention configurations, and structural perturbations. This validates the effectiveness of incorporating relational structures in enhancing the interpretability and precision of risk identification.

The study also highlights the sensitivity of model performance to key architectural parameters such as graph convolutional depth and attention head count. Careful tuning of these components enables the model to better extract meaningful features from heterogeneous graph data, while avoiding common issues such as overfitting and feature oversmoothing. Additionally, the model exhibits strong adaptability under moderate graph noise and class imbalance, making it suitable for real-world financial scenarios where credit data may be incomplete, noisy, or temporally dynamic. The findings contribute empirical evidence and methodological insights to the ongoing exploration of graph learning techniques in financial risk modeling.

Beyond the immediate task of credit risk classification, the methodology proposed in this paper can be extended to other domains where structured relationships and temporal dynamics play a key role. These include fraud detection, supply chain risk analysis, compliance monitoring, and macroeconomic network modeling. The ability to model inter-entity dependencies in a scalable and automated way opens up new possibilities for intelligent decision support systems in finance and beyond. The graph-based paradigm also provides a foundation for building more transparent, explainable, and adaptive models for high-stakes regulatory applications.

7. Future Work

Future work will focus on enhancing the temporal modeling capabilities of the framework by integrating dynamic graph neural networks and sequential learning mechanisms. Another promising direction is the incorporation of self-supervised or contrastive learning strategies to improve performance in low-label or noisy-label environments. Furthermore, developing more sophisticated graph augmentation and repair mechanisms will be essential for boosting model resilience in the face of missing or corrupted relational data. These advancements are expected to further improve the practical deployment of graph-based credit risk systems and support the development of more resilient and intelligent financial infrastructures.

References

[1] Wang D, Wu Z, Zhu B. Controlling shareholder characteristics and corporate debt default risk: Evidence based on machine learning[J]. Emerging Markets Finance and Trade, 2022, 58(12): 3324-3339.

[2] Olson L M, Qi M, Zhang X, et al. Machine learning loss given default for corporate debt[J]. Journal of Empirical Finance, 2021, 64: 144-159.

[3] Alam N, Gao J, Jones S. Corporate failure prediction: An evaluation of deep learning vs discrete hazard models[J]. Journal of International Financial Markets, Institutions and Money, 2021, 75: 101455.

[4] Xia S, Zhu Y, Zheng S, et al. A Deep Learning-based Model for P2P Microloan Default Risk Prediction[J]. Spectrum of Research, 2024, 4(2).

[5] Alonso Robisco A, Carbó Martínez J M. Measuring the model risk-adjusted performance of machine learning algorithms in credit default prediction[J]. Financial Innovation, 2022, 8(1): 70.

[6] Bussmann N, Giudici P, Marinelli D, et al. Explainable machine learning in credit risk management[J]. Computational Economics, 2021, 57(1): 203-216.

[7] Alonso A, Carbó J M. Understanding the performance of machine learning models to predict credit default: a novel approach for supervisory evaluation[J]. 2021.

[8] Alarfaj F K, Malik I, Khan H U, et al. Credit card fraud detection using state-of-the-art machine learning and deep learning algorithms[J]. Ieee Access, 2022, 10: 39700-39715.

[9] Wei S, Egressy B, Chen X, et al. Graph Dimension Attention Networks for Enterprise Credit Assessment[J]. arXiv preprint arXiv:2407.11615, 2024.

[10] Wang D, Zhang Z, Zhou J, et al. Temporal-aware graph neural network for credit risk prediction[C]//Proceedings of the 2021 SIAM International Conference on Data Mining (SDM). Society for Industrial and Applied Mathematics, 2021: 702-710.

[11] Feng B, Xu H, Xue W, et al. Every corporation owns its structure: Corporate credit rating via graph neural networks[C]//Chinese Conference on Pattern Recognition and Computer Vision (PRCV). Cham: Springer International Publishing, 2022: 688-699.

[12] Shi S, Luo W, Tse R, et al. SparseGraphSage: A Graph Neural Network Approach for Corporate Credit Rating[C]//Proceedings of the 2024 13th International Conference on Software and Computer Applications. 2024: 124-129.