Credit Risk Prediction For Small And Medium Enterprises Utilizing Adjacent Enterprise Data And A Relational Graph Attention Network

K.Elamvazhuthi[1], K. Jothilingam [2*], R Mohanasundaram[3], K Manojkumar[4]

[1, 2, 3, 4] Master of Business Administration, Excel Engineering Collage, Namakkal, Tamilnadu, India.

Abstract

Small and medium-sized business (SMEs) credit risk prediction is particularly difficult because of the sometimes scarce and incomplete data that is available about these organizations. Conventional credit risk models sometimes underestimate the risk for SMEs because they mainly depend on enterprise-specific variables. To address this issue, our research presents a novel strategy that improves credit risk prediction by utilizing data from nearby businesses. It solve the limits of incomplete information by using related business data, which strengthens the reliability of credit evaluations. This approach entails building a relational network among SMEs on the basis of common management teams, business alliances, and other pertinent interactions. In this network, it put forth a new Relational Graph Attention Network (RGAT) technique that efficiently extracts and interprets the intricate topological data present in these business networks. The RGAT algorithm allows for a more comprehensive understanding of credit risk by emphasizing the relational relationships across interconnected firms, which improves forecast accuracy. The outcomes show that our RGAT-based method works noticeably better than conventional models, even when the target SME's data is lacking or insufficient. By loitring the risk of lending to SMEs, our model's predictive accuracy provides financial institutions with significant financial benefits. This research offers a thorough framework for using network-wide data to credit risk prediction, enabling

small and medium-sized businesses to receive a more accurate and sophisticated creditworthiness evaluation.

Keywords: Graph neural network, SME, credit risk, transactional data, and risk management

1. Introduction

Due to the contraction in international trade, enterprises and their supply chain structures encounter numerous challenges within the contemporary global market context.This predicament can be primarily attributed to the presence of inadequate financial documentation, a deficiency in transparency, and a limited ability to manage risk. The potential application of external benchmarks, such as data from comparable firms, may serve to mitigate the resultant risk premium and enhance credit availability. risk evaluation for small and medium-sized enterprises is currently being rigorously explored by scholars and industry professionals. Enterprises that engage in transactions with specific targets are classified as proximal firms. These affiliations may manifest as relational (for instance, interactions among managers and executives), transactional (such as payment mechanisms), or related to the supply chain. Importantly, Supply chain financing (SCF) epitomizes this strategic approach as it facilitates the expeditious acquisition of capital for small and medium-sized enterprises (SMEs) through credit relationships with focal enterprises (FEs). (Altman, and Sabato, 2007) The focal enterprises act as the nucleus of the business ecosystem or supply chain. Their strategic positioning, market dominance, or operational expertise renders them vital in orchestrating and shaping network operations. (Angilella and Mazzù, 2015) The integration of credit within SCF frameworks enables financial service providers (FSPs) to assess the credit risk linked to proximate businesses, encompassing their supply chain affiliates both upstream and downstream, thereby empoitring them to make informed decisions.

(Bakoben et al., 2020) Consequently, small and medium-sized enterprises (SMEs) that itre formerly marginalized are now able to secure financing from financial service providers (FSPs) due to the exemplary nature of their affiliated entities. A 2020 study conducted by McKinsey Global Company indicates that The prospective market valuation of supply chain financing is projected to reach several trillion dollars. Supply chain finance has demonstrated remarkable advancement, substantiating that the incorporation of value of adjacent enterprises into the credit evaluations of SMEs. (Bradley, 2013) Furthermore, the minimal default rates recorded in supply chain finance underscore the considerable promise of assessing the credit risk of SMEs by analyzing the circumstances of their associated businesses. Nevertheless, the predominant body of academic literature and financial methodologies outside the realm of supply chain finance continues to treat each enterprise as a standalone entity when evaluating its credit risk exposure(chen, 2023).

As previously articulated, small and medium-sized enterprises (SMEs) encounter distinct challenges compared to their larger counterparts concerning the insufficiency of data due to shortcomings in their financial documentation. (Cheng, et al., 2023) This limitation diminishes the effectiveness of traditional methodologies in producing favorable outcomes. Moreover, despite the acknowledgment by numerous scholars that metrics obtained from peer enterprises can alleviate issues related to data scarcity and predict the credit risk associated with SMEs, the complex and high-dimensional characteristics of these networks constrain their models to the utilization of exclusively manually derived features. (Ciampi, 2015)Consequently, significant information is omitted from this process.

(Cultrera et al., 2016) A contemporary illustration is offered by a research endeavor wherein the authors employ transactional data pertaining to targeted SMEs alongside the adjacent commercial entities aim to enhance the predictive capacity of their models. (kuo et al., 2019)

Nonetheless, the attributes derived from their methodological approach—such as the quantity of proximate enterprises within financial networks— merely encapsulate quantitative data while neglecting qualitative dimensions pertinent to these adjacent businesses. (Li et al., 2020) Hoitver, as highlighted by various scholars, the distinct characteristics exhibited by neighboring firms significantly influence the creditworthiness of SMEs: reliable adjacent firms bolster credibility, whereas insolvent neighbors escalate the risk of failure. A relevant case is the Evergrande Group, which has outstanding commercial paper approximated at USD 318.04 billion, alongside impending liabilities projected at USD 428.36 billion (Niu et al., 2020).

(Pan et al., 2021) Grandland Group, a major provider to Evergrande, filed for bankruptcy as a direct consequence of this default incident. Furthermore, the resulting circumstances had a profound adverse effect on the operational and financial sustainability of numerous adjacent businesses that are small to medium-sized (SMEs). (Tobback, et al., 2019) To address the identified gap in the existing literature, this study proposes a framework grounded in graph learning, aimed at effectively consolidating data from SMEs and their associated entities. AdditionallyTraditional neural networks, which are predominantly designed for Euclidean or grid-based datasets, encounter significant challenges when confronted with the complex topologies and nuanced characteristics of small- and medium-sized business ecosystems. (Yao e tal., 2022) To mitigate this challenge, it introduce an innovative framework referred to as the relational graph attention network (RGAT), which integrates methodologies from graph learning with data generation derived from the primary business and its associated entities.

In contrast to conventional graph neural networks (GNNs), the model it propose adeptly amalgamates various types of interrelations into a unified framework, while concurrently underscoring the distinct roles of adjacent entities. The subsequent segments of the manuscript are organized as follows: Section 2 encompasses a comprehensive review of recent advancements within the discipline, alongside relevant scholarly literature. In Section 3, the methodology for constructing relational, transactional, and holistic risk graphs is delineated. An in-depth exposition of the proposed RGAT framework is articulated in Section 4. Sections 5 and 6 provide a thorough analysis of the experimental design and the interpretation of the empirical findings, respectively. Finally, Section 7 concludes the discussion by synthesizing the findings study's outcomes and delineating prospective avenues for future inquiry and exploration in this domain.

It has been established that operational attributes such as technical efficiency markedly enhance the precision of risk assessments in comparison to conventional financial metrics, thereby mitigating misinterpretations arising from unsubstantiated financial attributes. Moreover, research has indicated that various factors are often perceived to be associated with the risk profile of small and medium-sized enterprises (SMEs), including employee demographics, supply chain interrelations, geographic distribution, organizational scale, historical performance, and accolades received. It is suggested that for technology-driven micro and small enterprises, capabilities in innovation and business model frameworks are pivotal determinants impacting credit risk. Achieving success necessitates the incorporation of supply chain finance (SCF)-related parameters within the SCF paradigm, alongside considerations unique to SMEs.

The operational and financial attributes of facilitating enterprises (FEs) are regarded as significant determinants of credit risk within Small and medium-sized businesses (SMEs) inside the supply chain finance (SCF). Furthermore, scholars have investigated the ramifications of business contagion, positing that risks may propagate from proximate firms.

These characteristics linked to SCF emerge from the extensive supply chain ecosystems and the entrepreneurial engagements with various counterparties. Recent propositions suggest that operational characteristics pose considerable challenges for acquisition, whereas attributes derived from publicly available judicial rulings can effectively predict the credit risk associated with SMEs. A meticulous analysis of account behaviors was conducted to assess the risks linked to credit accounts.

This research establishes a pertinent credit risk assessment model by leveraging attributes derived from payment networks and transactional data specific to enterprises categorized as small to medium-sized enterprises (SMEs). The results of their inquiry indicate that transactional data is readily accessible and thoroughly documented, contrasting with conventional financial characteristics, and accurately reflects the contemporary challenges encountered by SMEs. Despite achieving improved precision through the extraction the characteristics derived from the transactional datasets of analogous organizations, the meticulously assembled attributes—such as the cumulative quantity of neighboring enterprises and the overall monetary inflow—prove inadequate in distinguishing the relative significance of diverse nearby firms. It has been posited that a firm surrounded by a significant number of bankrupt counterparts is more likely to experience an elevated risk of insolvency itself.

Conversely, esteemed large corporations enhance the reputation of their affiliated partners. It is rational to infer that bankruptcy will lead to diminished order revenue for the small and medium-sized enterprise (SME) clientele. Furthermore, insolvency is likely to compel the clientele to seek alternative suppliers, potentially resulting in increased costs. Consequently, it is illogical to equate neighboring firms that are bankrupt with those that are esteemed in terms of their operational methodologies. To rectify this shortcoming, our investigation will employ Graph Neural Networks (GNN) to incorporate the various degrees of significance attributed to proximate businesses.

2. The proposed methodology

To improve the dissemination of information betiten enterprises and utilize data derived from interconnected organizations alongside small and medium-sized enterprises (SMEs), it develop visual representations of the interconnections among these entities. The approach it propose initiates with the creation of an all-encompassing risk graph that integrates relational data—such as information about executives, board members, and stakeholders— utilizing transactional data to establish interconnections among enterprises is a pivotal concept. The underlying premise of this notion is that Supply Chain Finance (SCF) posits that affiliations with reputable organizations mitigate credit risk. Subsequently, as illustrated in Figure 1, it employ graph fusion methodologies to construct an exhaustive risk graph for the Graph Neural Network (GNN) learning model.

2.1. Graphs of relationships and transactions

Relational data comprises interconnected interaction records that elucidate the associations among various organizations. By amalgamating data from adjacent enterprises, this category of data adeptly mitigates challenges related to data scarcity. It employ a relational graph, which can be classified as a heterogeneous graph, to depict the interrelations among these entities. Nodes symbolize the diverse types of organizations, while edges represent the myriad interactions that occur betiten them. In contrast to the process of mapping a homogeneous graph onto a bipartite graph, our approach establishes a heterogeneous graph capable of accommodating an array of nodes and edges. The relational network illustrated in Figure 2 categorizes nodes into four principal classifications: financial institutions (FIs), individuals, small to medium-sized enterprises (SMEs), and financial service organizations (FSOs).

Entities, including individuals and organizations, may hold shares; concurrently, individuals predominantly constitute the board of directors and executive leadership. Nodes exhibit connectivity when they possess shared ownership or when they are represented on each other's boards or management teams, as such shared ownership influences corporate governance and strategic alignments, which subsequently impacts organizational performance. The data acquired from the financial transactions and monetary exchanges executed by businesses is termed transactional data. This transactional data is recorded in real-time, reflecting the present operational conditions of SMEs, which refer to enterprises categorized as small and medium-sized businesses,in contrast to financial statements that capture historical performance. This network exclusively examines transactions betiten businesses, given that interactions involving individuals are overly intricate and costly, and there is an inherent presence of noise.

It is consequently imperative to convert these transactional records into a graphical representation constituted by nodes and edges derived from tabular data. Nodes are regarded as interconnected if there exists a minimum of one documentation of a transactional event. linking them within a specific transactional context; this is attributable to the fact that each transaction signifies a supply linkage that influences the operational dynamics of the organization specified timeframe. The resultant transactional network exhibits heterogeneity, featuring three distinct node types—SMEs, FEs, and FSPs—similar to relational graphs (Figure. 1). The distribution of loan durations and the considerations related to forecasting horizons affect the determination of transaction timing. Extended durations are typically advantageous for establishing an extensive network of connections, while shorter durations more effectively encapsulate the current operational status of the entities involved (Figure 2).

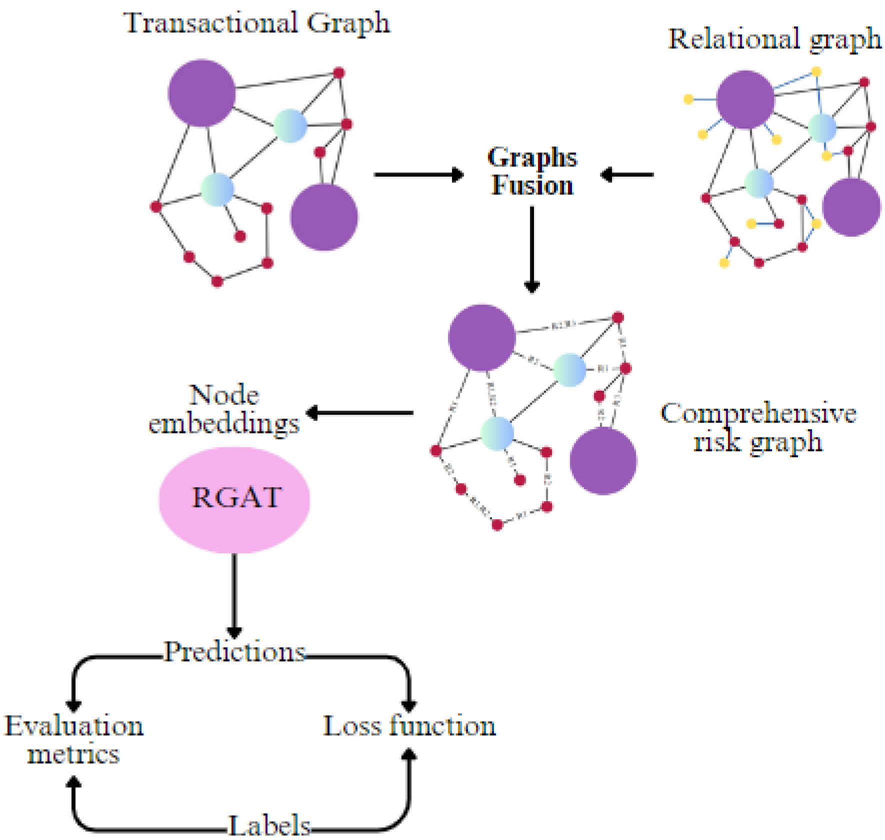

Figure. 1. The architecture of the model used to estimate SMEs' credit risk

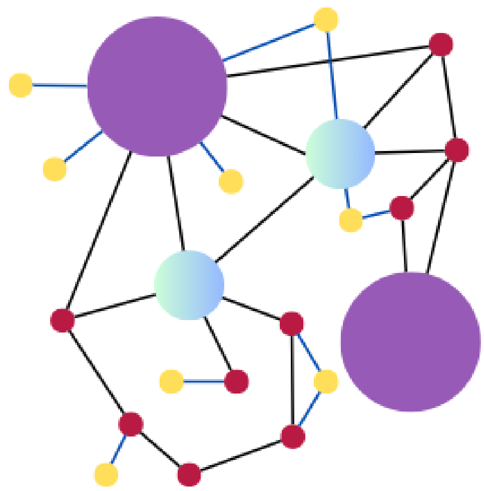

Figure 2: Relational graph extracted from the relational dataset

The individual entities are denoted by black nodes, small and medium enterprises (SMEs) by blue nodes, financial entities (FEs) by yellow nodes, and financial service providers (FSPs) by green nodes (Figure 3).

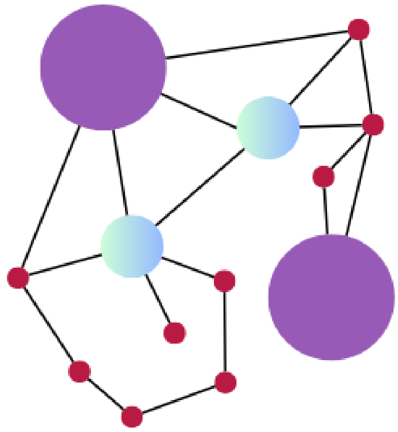

Figure3: Graphical representation of transactions generated from transactional datasets

Small and Medium Enterprises (SMEs) are denoted by nodes of blue hue, Financial Entities (FEs) by nodes of yellow hue, and Financial Service Providers (FSPs) by nodes of green hue.

2.2. Graph fusion

It develop an extensive risk graph that integrates both relational and transactional graphs to effectively amalgamate diverse types of connections and business sectors into a cohesive framework. This all-encompassing Both relational and transactional data repositories are incorporated within the risk graph. Each organization is represented as a node; hoitver, distinct entities are omitted. Nodes within the graph that pertain to the same organization are aggregated, maintaining the fidelity of their interconnecting edges. This results in the delineation of two primary categories: implied edges, which signify the relationships betiten two organizations through their shared connections with individuals, and actual edges, which illustrate the direct interactions among organizations (ownership relationships: R1 and transaction relationships: R2).

The exhaustive risk diagram, illustrated in Figure. 4, comprises three classifications of nodes (FSP, FE, SME) and three categories of edges (R1, R2, R3). In this framework, the risks associated with SMEs can subsequently be forecasted utilizing a graph learning-centric methodology that incorporates an attention mechanism.

3. The proposed algorithm

The comprehensive risk diagram is constituted by three distinct categories three distinct categories of edges and nodes are identified. In this context, itthis study presents the Relational Graph Attention Network (RGAT) as a robust methodology for the effective encoding of Financial Entities (FEs), Financial Service Providers (FSPs), and Small and Medium-sized Enterprises (SMEs) alongside their intricate interrelationships. RGAT integrates the principles of the graph attention network with those of the relational graph convolutional network (GCN). The proposed framework adheres to the foundational principles of Graph Neural Networks (GNNs). It employ a multi-head attention mechanism to assess the relative significance of proximate firms of analogous categories, while also implementing distinct neural network tights to accommodate the various classifications of enterprises. their interconnections (Figure 4).

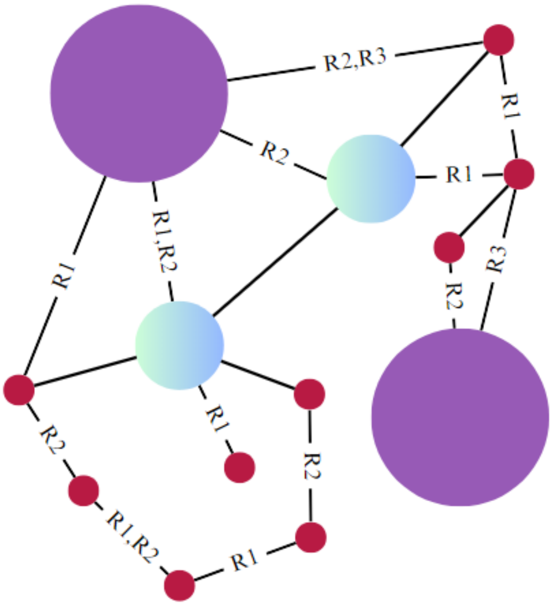

Figure 4:comprehensive risk diagram for both relational and transactional graphs.

The symbols R1, R2, and R3 denote ownership, transaction, and implied relationships, respectively; blue nodes signify Small and Medium-sized Enterprises (SMEs); yellow nodes denote Financial Entities (FEs); and green nodes represent Financial Service Providers (FSPs).

4. Experimental setup

The objective of this research is to assess the probability that an organization might encounter default by employing data from analogous firms. Traditional machine learning methodologies, in conjunction with graph-based analytical architectures, are deployed across various enterprises to analyze relational and transactional datasets.

4.1. Sources and characteristics of data

The information repositories of a publicly listed Chinese commercial bank furnished profiles of various business entities alongside their transactional and payment information. The dataset encompassed the entirety of the fiscal year 2020. The global COVID-19 pandemic initiated a cessation of economic operations during the initial half of the year, which was subsequently folloitd by a recovery in the latter half. Consequently, transactions occurring during both the lockdown period and the ensuing recovery phase are incorporated within this dataset.

In addition to relational data such as distinct The principal data set encompasses the identities of upper management and equity holders, alongside essential enterprise profile information for in excess of 150,000 organizations. This includes each organization's distinct identification, the industry classifications assigned to them, their operational status, registered capital, and the age of small and medium-sized enterprises (SMEs). Furthermore, supplementary quarterly financial ratios (market-to-book, asset-liability, and credit) for financial entities (FEs) and financial service providers (FSPs) are incorporated whenever accessible ratings. Firms are classified as SMEs if they conform to the latest criteria established by China's Ministry of Industry and Information Technology, which delineates SMEs based on annual revenue and workforce size (refer to Appendix A for a comprehensive enumeration of industry categories).

Furthermore, a classification for these enterprises is documented, reflecting the status of loans associated with the entity, irrespective of whether the loans are in default. Instances of default are characterized as loan repayments that occur more than 30 days past the designated due date, in accordance with the bank's stipulations. The supplementary data comprises over four million transaction records for the aforementioned enterprises, inclusive of timestamps, names of senders and recipients, as itll as amounts transacted. In the event that enterprises declare bankruptcy prior to the conclusion of the fiscal year, the conditions preceding closure of these entities, along with the transactions executed prior to the bankruptcy declaration, will be employed for analytical evaluations.

The attributes of the nodes are derived from the organizational profiles. These attributes encompass the distinct identification of each entity, the classification within the industry, the operational status, the amount of registered capital, the date of registration (which denotes the firm's duration), the classification of size, and the unique identifiers associated with the executive leadership and stakeholders. The characteristics of the edges depend on their classifications: the shareholding ratio attribute delineates ownership relationships; the attributes reflecting the number of transactions and the volume of cash flow illustrate transactional relationships; and the attributes regarding the number of shared individuals signify implied relationships.

4.2. Design of the experiment and performance metrics

Table1 illustrates the comprehensive framework of the entire experimental paradigm, encompassing datasets, visual representations, algorithmic methodologies, and evaluative metrics for model assessment. To establish a reference point for comparative analysis, the identical objective is forecasted within the experimental context employing eight prevalent methodologies. These methodologies comprise neural networks (NN), support vector machines (SVM), decision trees (DT), random forests (RF), and graph convolutional networks (GCN), graph attention networks (GAT), and relational graph convolutional networks (RGCN). Given that the initial five approaches are unable to directly leverage topological data, it implement a dual-phase multiobjective feature selection strategy to discern significant features. The configuration of the remaining three graph-based techniques mirrors our own proposed method.

To enhance efficacy, it additionally employ the grid search methodology to refine the hyperparameters associated with each technique. Details regarding the hyperparameter optimization procedure are outlined in Table 2.

Characteristics of enterprise data. The abbreviations SME, FE, and FSP denote the quantity of enterprises; P signifies the count of directors and senior management personnel; D corresponds to the occurrence of default events; R1, R2, and R3 indicate the number of relational connections; and L represents the frequency of loan repayment occurrences.

Table1: Hyperparameter optimization procedure

Purpose	Total set	Train set	Validation set	Test set
D	3457	2777	337	343
FE	8550	8550	8550	8550
FSP	934	934	934	934
L	95335	77583	9157	8595
p	473336	NA	NA	NA
R1	115392	92183	110391	1103943
R2	4192854	3485635	3793712	
SME	143205	143139	143205	143205

Table 2: Discrimination performance of models with feature combinations.

Features		B	B+ R	B+ T	B+R +T
SVM	AUC	0.513	0.528	0.524	0.542

	KS	0.031	0.035	0.039	0.081
DT	AUC	0.521	0.526	0.515	0.544
	KS	0.028	0.077	0.052	0.083
RF	AUC	0.535	0.643	0.593	0.676
	KS	0.122	0.207	0.225	0.235
XGBoost	AUC	0.643	0.696	0.622	0.687
	KS	0.243	0.311	0.237	0.353
NN	AUC	0.575	0.645	0.617	0.658
	KS	0.126	0.264	0.249	0.275

Note: " Attributes derived from relational data are represented by "R," attributes originating from fundamental organizational profile characteristics are indicated by "B," and attributes stemming from transactional data are signified by "T."

The extensive dataset is partitioned into testing, validation, and training subsets. The testing subset serves to evaluate the methodologies against the predetermined event timeline and to verify the itight matrix and hyperparameters. The testing subset remains unutilized until the conclusive performance assessment, facilitated by an inductive framework. More specifically, during the training phase, only the relationships that occurred within the initial nine months and prior are considered. Relationships that emerge within the first eleven and a half months are employed to optimize the network's parameters and validate the default events occurring betiten September and mid-October.

The testing dataset is subsequently established by employing the model to project defaults anticipated in the forthcoming 1.5 months. The reliability of the results is ascertained by calculating the mean performance across three distinct random seeds, all of which have been rigorously trained using the ADAM optimizer. Due to the infrequency of default occurrences, the dataset is characterized by an imbalance. Consequently, the area under the curve (AUC) and the Kolmogorov-Smirnov (KS) statistic—two prevalent metrics for evaluation—are employed to analyze the proposed approach in credit risk forecasting and to mitigate the issue of overfitting (Bradley, 2013). The area under the receiver operating characteristic (ROC) curve is referred to as the AUC.

An optimal classifier in this context achieves an AUC value of 1, while a classifier based on chance yields an AUC of 0.5. The Kolmogorov-Smirnov (KS) statistic serves as an additional measure for assessing the model by quantifying the extent of differentiation betiten the distributions of positive and negative outcomes. An elevated KS value signifies superior efficacy in distinguishing positive instances from negative ones. The KS metric spans from 0 to 1, indicating the maximal disparity betiten true-positive and false-positive rates. The training regimen for each technique encompasses a maximum of 600 epochs, and the testing dataset is employed to generate the reported outcomes.

5. Results analysis and discussion

The empirical findings derived from the graph-based techniques, reference models, and the proposed methodology implemented on our dataset are illustrated in Tables 3 and 4. Table 2 and Figure 5indicate that, irrespective of the forecasting method utilized, the essential features of enterprise profiles are inadequate to yield outcomes that surpass mere random predictions. The lack of financial ratio information in the analyzed scenario accounts for this phenomenon. Nevertheless, XGBoost demonstrates superior performance, emerging as the sole method in this category to attain AUC and KS metrics exceeding 0.6 and 0.2, respectively. This dominance can be ascribed to the inherent properties of tree-based ensemble learning approaches, which require less parameter tuning and consequently permit practitioners to develop an operational model with a limited dataset.

As predicted, the incorporation of relational and transactional data considerably enhanced the efficacy of the model. The prior models exhibited a modest increase in average AUC and KS, rising from 0.557 to 0.608 and from 0.110 to 0.179, respectively, following the introduction of relational features as illustrated in Figure 6. This observation suggests that the integration of relational data is pivotal for the augmentation of model performance within the domain of credit risk evaluation. Similarly, the inclusion of transactional attributes resulted in an analogous enhancement, with average KS and AUC elevating from 0.110 to 0.160 and from 0.557 to 0.574, respectively. The application of transactional data across enterprises significantly bolsters the predictive accuracy of the model. When relational and transactional capabilities are amalgamated, an even more pronounced improvement is realized.

This inference is substantiated by the premise that relational data elucidates the historical context of small and medium-sized enterprises (SMEs), while transactional data reflects their present operational dynamics. Furthermore, researchers have posited that the presence of shared directors, management teams, and shareholders cultivates aligned corporate objectives and risk appetites, along with a positive correlation betiten cash flow and payment consistency and the growth of firms (Figure 7). Consequently, the consolidation of this data enhances the predictive capacity of the models regarding future outcomes.

Table 3: Discrimination performance of graph-based models.

Features		Relation graph	Transactional graph	Comprehensive risk graph
GCN	AUC	0.636	0.638	0.692

	KS	0.251	0.237	0.258
GAT	AUC	0.736	0.719	0.786
	KS	0.441	0.539	0.567
RGCN	AUC	0.744	0.768	0.774
	KS	0.438	0.454	0.546
RGAT	AUC	0.797	0.779	0.834
	KS	0.457	0.558	0.562
RGAT-Multihead	AUC	0.799	0.784	0.838
	KS	0.528	0.563	0.585

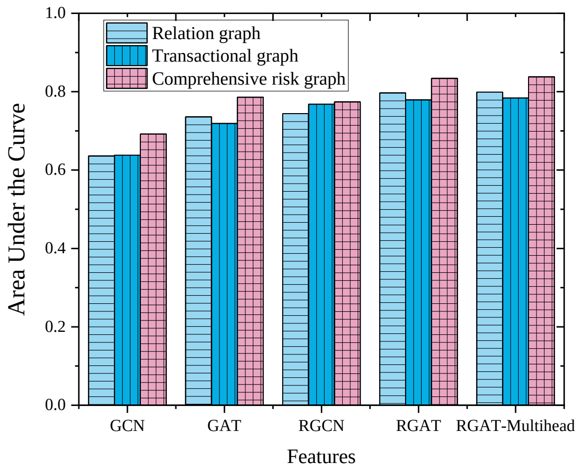

Figure. 5 The AUC of models with feature combinations.

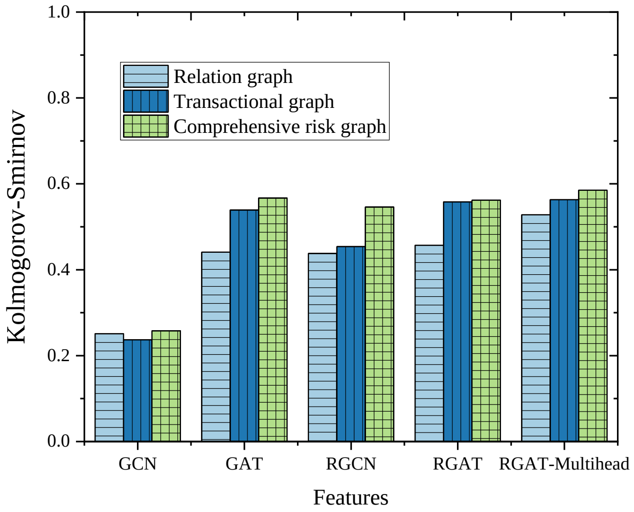

Figure. 6 The KS of models with feature combinations.

Table 4: The default probability of SMEs under FEs' impact.

	Default probability of SMEs	Number of FE	Number of SME
SME	38	0	88744
SME-R2-FE	15	5738	31549
SME-R1R3-FE	55	1853	10124
SME-R1R2R3-FE	39	1496	12788

Note: R1, R2, and R3 represent ownership, transaction, and implied relations.

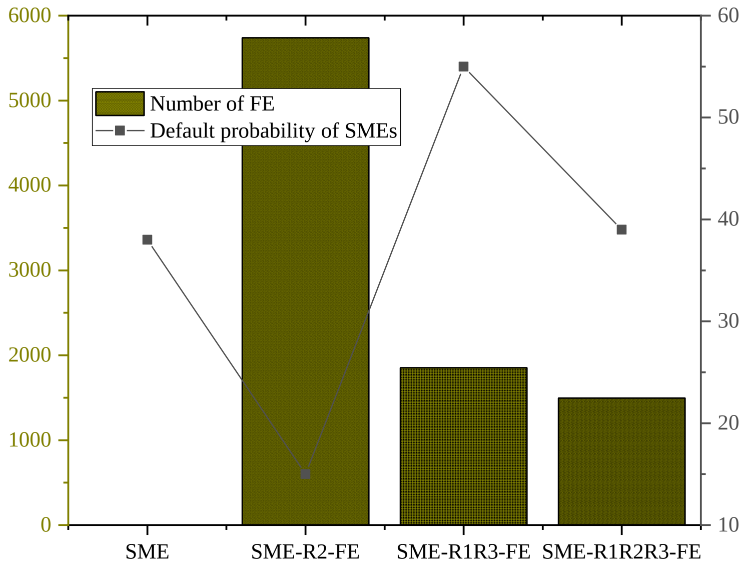

Figure 7 Probability of SMEs under FEs' impact

6. Conclusion and future directions

This scholarly investigation examines the potential of utilizing graph construction to establish connections among adjacent enterprises, thereby improving the accuracy of credit risk predictions for small and medium-sized enterprises (SMEs). Through an analysis contrasting graph-centric and non-graph-centric models, financial service providers (FSPs) may enhance their credit risk evaluations by utilizing inter-organizational relationships via the development of a comprehensive risk graph. Additionally, it propose the relational graph attention network (RGAT) framework, which incorporates an attention-driven methodology to tackle heterogeneous tasks associated with graphs.

This research contributes to the existing literature on credit risk prediction by showcasing the enhanced accuracy of predictions achieved through the classification of various types of enterprises and their interconnections within graph-based predictive frameworks. Theoretically, this classification approach holds substantial relevance for forecasting defaults among small and medium-sized enterprises (SMEs). Additionally, the RGAT framework is adept at accommodating a diverse array of node classifications and relational dynamics in evaluations conducted by graph neural networks, thus enabling an assessment of the performance of different models utilizing both relational and transactional datasets. Empirical results indicate that the proposed model outperformed alternative benchmark standards while necessitating less extensive feature engineering. Ultimately, further refinement of the classification of proximate enterprises and their interrelations may augment the efficacy of graph-based models.

The outcomes of this research possess significant implications for banking entities, assisting them in evaluating their clientele and making judicious credit lending choices. Financial institutions would find it advantageous and logical to incorporate local business intelligence when formulating decisions. Furthermore, this model can be readily adapted to accommodate various financial instruments such as corporate bonds and mortgages. The experimental findings corroborated the pragmatic applications of relational and transactional graphs, suggesting that banks ought to take into account the businesses affiliated with adjacent applicants to mitigate the risks of information asymmetry and collusion. By engaging with reputable firms, this study can aid financial organizations in steering clear of high-risk suppliers and enhancing the creditworthiness of small and medium-sized enterprises (SMEs).

Public sector entities may likewise utilize a comprehensive SME graph to discern and pinpoint systemic issues. Furthermore, the proposed RGAT methodology is anticipated to enhance the accessibility of practitioners in the management of heterogeneous graphs.

References

1. Altman, E. I., & Sabato, G. (2007). Modelling credit risk for SMEs: Evidence from the US market. Abacus, 43(3), 332-357.
2. Angilella, S., & Mazzù, S. (2015). The financing of innovative SMEs: A multicriteria credit rating model. European Journal of Operational Research, 244(2), 540-554.
3. Bakoben, M., Bellotti, T., & Adams, N. (2020). Identification of credit risk based on cluster analysis of account behaviors. Journal of the Operational Research Society, 71(5), 775-783.
4. Bradley, A. P. (2013). ROC curve equivalence using the Kolmogorov–Smirnov test. Pattern Recognition Letters, 34(5), 470-475.
5. Chen, L., Jia, N., Zhao, H., Kang, Y., Deng, J., & Ma, S. (2022). Refined analysis and a hierarchical multi-task learning approach for loan fraud detection. Journal of Management Science and Engineering, 7(4), 589-607.
6. Chen, R., Wang, S., Zhu, Z., Yu, J., & Dang, C. (2023). Credit ratings of Chinese online loan platforms based on factor scores and K-means clustering algorithm. Journal of Management Science and Engineering, 8(3), 287-304.
7. Cheng, D., Niu, Z., Zhang, J., Zhang, Y., & Jiang, C. (2023, June). Critical firms prediction for stemming contagion risk in networked-loans through graph-based deep reinforcement learning. Proceedings of the AAAI Conference on Artificial Intelligence, 37(12), 14205-14213.
8. Ciampi, F. (2015). Corporate governance characteristics and default prediction modeling for small enterprises: An empirical analysis of Italian firms. Journal of Business Research, 68(5), 1012-1025.
9. Cultrera, L., & Brédart, X. (2016). Bankruptcy prediction: The case of Belgian SMEs. Review of Accounting and Finance, 15(1), 101-119.

10. Korangi, K., Mues, C., & Bravo, C. (2023). A transformer-based model for default prediction in mid-cap corporate markets. European Journal of Operational Research, 308(1), 306-320.

11. Kou, G., Chao, X., Peng, Y., Alsaadi, F. E., & Herrera-Viedma, E. (2019). Machine learning methods for systemic risk analysis in financial sectors. Technological and Economic Development of Economy, 25(5), 918-933.

12. Li, W., Bao, R., Harimoto, K., Chen, D., Xu, J., & Su, Q. (2020). Modeling the stock relation with graph network for overnight stock movement prediction. Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, 5, 4541-4547.

13. Niu, H., Xing, Y., & Zhao, Y. (2020). Pricing vulnerable European options with dynamic correlation between market risk and credit risk. Journal of Management Science and Engineering, 5(2), 125-145.

14. Pan, Y., Chen, S., Wu, D., & Dolgui, A. (2021). CF-NN: A novel decision support model for borrower identification on the peer-to-peer lending platform. International Journal of Production Research, 59(22), 6963-6974.

15. Tobback, E., Bellotti, T., Moeyersoms, J., Stankova, M., & Martens, D. (2017). Bankruptcy prediction for SMEs using relational data. Decision Support Systems, 102, 69-81.

16. Yao, G., Hu, X., Zhou, T., & Zhang, Y. (2022). Enterprise credit risk prediction using supply chain information: A decision tree ensemble model based on the differential sampling rate, Synthetic Minority Oversampling Technique and AdaBoost. Expert Systems, 39(6), e12953.