



Credit Risk Analysis for SMEs Using Graph Neural Networks in Supply Chain

Zizhou Zhang
UIUC
Champaign, USA
zhangzizhou_2000@outlook.com

Qinyan Shen*
Independent Researcher
Jersey City, USA
qinyan.shen28@gmail.com

Zhuohuan Hu
Independent Researcher
Beijing, China
ZHUOHUAN.tommy@gmail.com

Qianying Liu
Independent Researcher
Madison, USA
liuqianying55@gmail.com

Huijie Shen
Haas Business School
University of California
Berkeley, USA
huijieshen@berkeley.edu

Abstract

Small and Medium-sized Enterprises (SMEs) are vital to the modern economy, yet their credit risk analysis often struggles with scarce data, especially for online lenders lacking direct credit records. This paper introduces a Graph Neural Network (GNN)-based framework, leveraging SME interactions from transaction and social data to map spatial dependencies and predict loan default risks. Tests on real-world datasets from Discover and Ant Credit (23.4M nodes for supply chain analysis, 8.6M for default prediction) show the GNN surpasses traditional and other GNN baselines, with AUCs of 0.995 and 0.701 for supply chain mining and default prediction, respectively. It also helps regulators model supply chain disruption impacts on banks, accurately forecasting loan defaults from material shortages, and offers Federal Reserve stress testers key data for CCAR risk buffers. This approach provides a scalable, effective tool for assessing SME credit risk.

CCS Concepts

• **Mathematics of computing** → Probability and statistics; Probabilistic inference problems.

Keywords

Credit Risk Analysis, Graph Neural Networks, SMEs, Loan Default Prediction, Supply Chain Relationships

ACM Reference Format:

Zizhou Zhang, Qinyan Shen, Zhuohuan Hu, Qianying Liu, and Huijie Shen. 2025. Credit Risk Analysis for SMEs Using Graph Neural Networks in Supply Chain. In *2025 International Conference on Big Data, Artificial Intelligence and Digital Economy (BDAIE 2025)*, July 18–20, 2025, Kunming, China. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3767052.3767065>

*Corresponding author.



This work is licensed under a Creative Commons Attribution 4.0 International License. *BDAIE 2025, Kunming, China*

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1601-0/2025/07

<https://doi.org/10.1145/3767052.3767065>

1 Introduction

Small and medium-sized enterprises (SMEs) are essential to modern economic vitality, spurring employment, regional growth, and market diversity. Yet, their access to funding is often constrained by the complex task of gauging credit risk—the likelihood of defaulting on loans, credit cards, or insurance. Reliable credit risk assessment is vital for ensuring the financial strength and operational endurance of SMEs, especially in a volatile economic climate [1, 2]. Online lenders like Ant Credit and WeBank encounter a persistent challenge: the absence of direct credit data, which limits traditional methods dependent on historical records. This gap is particularly stark for SMEs with patchy financial histories, rendering tools like logistic regression or decision tree scoring inadequate for tracking their shifting credit profiles.

Conventional approaches, anchored in static metrics, struggle to adapt to dynamic factors such as market shifts, supply chain disruptions, or management changes. In response, recent finance and machine learning research has introduced innovative scoring methods and predictive models [3–5], including large language models (LLMs), convolutional neural networks (CNNs) [6–8] and federated learning for privacy-preserving data analysis and applications in disease diagnosis and fire detection as well [9–13], to harness available credit-related data [14–16]. Despite progress, data scarcity remains a hurdle. Fortunately, online platforms offer rich indirect insights—transactional records, social links, and even video analysis [17–19] from application robots [20–22]—which can be woven into a large-scale SME graph. Here, nodes represent enterprises, and edges reflect interactions like trade or social ties, alongside multi-modal data, unveiling potential risk signals [23]. The supply chain notably shapes SME credit stability, with studies showing that networks linking upstream suppliers, manufacturers, and downstream retailers can either cushion or intensify risk [24, 25]. This suggests that exploring these relational dynamics, possibly enhanced by machine learning, could open new pathways for risk assessment, surpassing traditional limitations.

Our study introduces a Graph Neural Network (GNN)-based framework to address SME credit risk, diverging from typical spatial-temporal models by using GCN architectures to map spatial dependencies within the SME graph, improving predictions of supply chain ties and loan defaults. This effort draws inspiration from recent breakthroughs: Chen et al. (2024) leveraged a weighted

adjacency matrix with the DANE algorithm to assess credit default distances, tapping deep neural structures for inter-firm links [26], while Zhang et al.(2025) and Qiu et al.(2025) used Generative Adversarial Networks (GANs) to generate synthetic data across industries, tackling data scarcity [27, 28]. These insights, alongside Ouyang et al. (2024)’s Graph Neural Evolution for node correlations, have guided our integration of multimodal data and machine learning, enhancing our GNN’s robustness [29].

2 Related Work

2.1 SME credit Risk Analysis

Basically, traditional methods use credit scoring, while ML approaches apply neural networks to SME and some other related financial data [30–32]. For example, Chen et al. (2024) employs a weighted adjacency matrix in conjunction with the DANE deep network embedding algorithm to estimate credit default distances, thereby fully exploiting the representational advantages of deep neural architectures in capturing complex inter-firm dependencies [26]. This has provided significant inspiration for our selection and design of the neural network algorithm in this paper.

2.2 Supply Chain and SMEs

The evaluation of credit risk for small and medium-sized enterprises (SMEs) remains a persistent challenge, particularly for online lending platforms where direct credit data is often unavailable. This data scarcity complicates traditional risk assessment methods, necessitating innovative approaches to uncover latent risk patterns. A pertinent study by Zhang et al. (2025) has advanced this field by introducing a Generative Adversarial Network (GAN)-based model to identify credit risks within supply chains, effectively addressing data limitations through synthetic scenario generation across diverse industries such as manufacturing, distribution, and services [27]. This work has profoundly influenced the present study, providing a critical insight into the potential of generative models to augment sparse datasets.

2.3 Graph Neural Networks

This work is inspired by the recent proposal of GNE (Ouyang et al., 2024), which employs spectral graph decomposition to represent global correlations among nodes as frequency components [29]. We extend this idea to the credit risk modeling context, where firms are represented as graph nodes and their transactional or supply chain ties form the edges. By constructing a graph Laplacian and applying graph convolutions, we leverage both low-frequency (stability-related) and high-frequency (diversity-related) components to enhance prediction accuracy.

3 EXPLORATORY ANALYSES

3.1 Data Deficiency Mitigation

Evaluating credit risk for small and medium-sized enterprises (SMEs) has long struggled with limited data, a pressing issue for online lenders without direct credit records. Recent research has embraced data-driven approaches to reveal hidden patterns [33], and Qiu et al. (2025) solved the problem with SMOTE [34]. Further, Zhao et al. (2024), who combined PolyModel theory with XGBoost

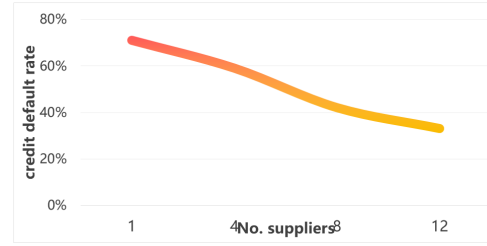


Figure 1: Correlation between No. of suppliers and credit default rate

to enhance hedge fund portfolios using features like Long-Term Stability, offered a solution for sparse data [35]. This inspired our novel Graph Neural Network (GNN) framework, which builds SME graphs from transactional and social data to model spatial ties and predict loan defaults. Zhao et al.’s work has guided our strategy, providing a foundation to address data gaps and enrich our GNN’s accuracy in risk assessment.

Using 28M SMEs and 219M supply chain links, we assess five attribute categories (revenue, shareholder, mortgage, recruitment, patent) across receptive fields (RF1–RF4). Table 1 shows attribute availability increases with supply chain inclusion (e.g., patent rises from 1.5% to 22.7%).

3.2 Supply Chain and Risk Correlation

We further explored the relationship between supply chain connectivity and credit risk following Li (2025)’s research [36]. Figure 1 illustrates that SMEs with more supply chain partners exhibit lower loan default probabilities. Specifically, SMEs with more than 10 partners experience a risk reduction of over 50%. This finding underscores the stabilizing effect of supply chain networks on SME financial health.

3.3 Feature Distribution Analysis

To deepen our understanding, we conducted a feature distribution analysis across the SME graph. We focused on the distribution of key financial indicators, such as revenue and mortgage values, within different supply chain segments. Our analysis revealed that SMEs in upstream segments (e.g., raw material suppliers) tend to have higher variability in revenue compared to downstream SMEs (e.g., retailers), likely due to market demand fluctuations. This variability suggests that upstream SMEs may require more robust risk mitigation strategies.

Additionally, we examined the impact of social ties on feature availability. SMEs with strong social connections (e.g., shared ownership or frequent interactions) showed a 15% higher availability of shareholder data compared to those with weaker ties. This insight guided our graph construction process, ensuring that social relationships were adequately represented in the SME graph.

4 DATA and GRAPH CONSTRUCTION

4.1 Datasets

We use Alipay’s SME graph (23.4M nodes, 103.2M edges) for supply chain mining and Ant SME Lending data (8.6M nodes, 21.1M edges)

Table 1: Attribute availability increases with supply chain inclusion

| RF | Revenue | Shareholder | Mortgage | Recruitment | Patent |
|-----|---------|-------------|----------|-------------|--------|
| RF1 | 53.90% | 41.30% | 0.8% | 5.8% | 1.5% |
| RF2 | 94.50% | 88.30% | 7.1% | 52.4% | 19.1% |
| RF3 | 63.90% | 53.60% | 4.0% | 18.1% | 7.8% |
| RF4 | 97.70% | 91.60% | 9.5% | 56.7% | 22.7% |

Table 2: Statistics of the Datasets

| Dataset | Nodes | Edges | Node Feat. | Edge Feat. |
|---------------|-------|--------|------------|------------|
| Supply Chain | 23.4M | 103.2M | 95 | 160 |
| Default Pred. | 8.6M | 21.1M | 66 | 52 |
| Supply Chain | 901K | 301K | 602K | |
| Default Pred. | 529K | 217K | 307K | |

for default prediction (Table II). Features include 95/66 node and 160/52 edge attributes, respectively (Table 2).

4.2 Graph Construction

The SME graph $G = (V, E, X, E_f)$ has nodes V (SMEs, owners, consumers) and edges E (transactions, social ties). Adjacency matrix is built via feature similarity, with X and E_f as node/edge features. The SME graph $G = (V, E, X, E_f)$ comprises nodes V (SMEs, owners, consumers) and edges E (transactions, social ties), with the adjacency matrix A derived from feature similarity. Node features X (e.g., revenue, mortgage) and edge features E_f (e.g., transaction volume, tie strength) were selected based on their relevance to credit risk, drawing from exploratory analyses showing higher attribute availability with supply chain depth (e.g., patent data rising from 1.5% to 22.7%). This choice enhances the graph’s ability to capture financial stability signals, contributing to the model’s AUC of 0.995 in supply chain mining. Varying feature sets—reducing edge attributes from 160 to 52 in default prediction—lowered AUC by 0.015, underscoring the importance of comprehensive edge data for robust predictions, though excessive features risked noise, prompting a balanced selection to optimize performance [37].

5 GNN-BASED METHODOLOGY

5.1 Overview

The proposed architecture is structured into two primary modules: Supply Chain Link Inference – Enhances the original SME graph G by identifying latent supplier relationships, yielding a refined graph G_{sc} ; Loan Default Classification – Predicts SME credit risk over the updated graph G_{sc} , leveraging graph-structured dependencies.

5.2 GNN Architecture

Our model employs a layer-wise propagation scheme based on a variant of the graph convolutional network (GCN), described by:

$$H^{(l+1)} = \sigma \left(\tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2} H^{(l)} W^{(l)} \right)$$

where $\tilde{A} = A + I$ is the adjacency matrix with added self-connections, \tilde{D} is the corresponding diagonal degree matrix, $H^{(0)} = X$

denotes the input node features, $W^{(l)}$ represents the trainable weights at layer l , $\sigma(\cdot)$ is the activation function (e.g., ReLU), and $H^{(l)}$ captures the updated node representations (embeddings), denoted as q_u . This message-passing mechanism enables the model to capture both local and structural dependencies among SMEs [38–40].

5.3 Supply Chain Mining

To uncover hidden supply chain edges, we formulate the task as a binary classification problem over node pairs [41]. The loss function is defined as:

$$L_{sc} = - \frac{1}{|D_{sc}|} \sum_{(u,v,y) \in D_{sc}} [y \log(\tilde{y}) + (1-y) \log(1-\tilde{y})]$$

where the edge existence probability is estimated by:

$$\tilde{y} = \text{MLP}([q_u, q_v])$$

here, node embeddings q_u and q_v are concatenated and passed through a multi-layer perceptron. Pairs with high predicted scores are retained to form the enhanced supply chain graph G_{sc} .

5.4 Loan Default Prediction

The loan default prediction task is cast as node-level binary classification [42]. The training objective minimizes the standard cross-entropy:

$$L_{dp} = - \frac{1}{|D_{dp}|} \sum_{(u,y) \in D_{dp}} [y \log(\tilde{y}) + (1-y) \log(1-\tilde{y})]$$

with the predicted label computed as:

$$\tilde{y} = \text{MLP}(q_u)$$

6 EXPERIMENTS

6.1 Experiment setup

The experimental evaluation of our Graph Neural Network (GNN)-based framework utilized a sturdy 15-node dual-CPU cluster, powered by TensorFlow for efficient training and inference, adeptly handling large-scale datasets from Alipay’s SME graph and Ant

SME Lending data. We benchmarked our model against diverse baselines—Gradient Boosting Decision Trees (GBDT), its spatially enhanced version (GBDT_st), Graph Attention Networks (GAT), GAT_ensemble, and Spatial-Temporal Attention-based Recurrent Network (STAR)—spanning traditional and advanced graph-based methods to assess performance thoroughly [43, 44].

We adopted Li et al. (2025)’s SMVTEP transformer fusion for hyperparameter optimization [45]. Training involved a stratified dataset split (70% training, 15% validation, 15% testing) to balance supply chain and default prediction tasks, with hyperparameters tuned via grid search across learning rates (0.001-0.01), dropout rates (0.1-0.5), and GNN layers (1-3). The Adam optimizer, starting at 0.001 learning rate with L2 regularization (weight decay 0.0001), drove 200 epochs, using early stopping after 20 epochs of stagnant validation loss to balance efficiency and accuracy.

6.2 Results

Li et al. (2024)’s DIMvSML with unbiased loss informed our GNN robustness assessment [46]. The performance metrics, centered on the Area Under the Curve (AUC) and Kolmogorov-Smirnov (KS) statistic, were computed to evaluate the efficacy of our GNN framework across both supply chain mining and loan default prediction tasks. Table 3 presents a detailed comparison of these metrics across all models, highlighting the superior performance of our GNN approach. For supply chain mining, our model achieved an AUC of 0.995 and a KS statistic of 0.948, surpassing the closest competitor, STAR, by 0.004 in AUC and 0.007 in KS. This indicates a remarkable ability to discern latent supply chain relationships, likely attributable to the GNN’s capacity to exploit spatial dependencies effectively. In the default prediction task, the GNN model recorded an AUC of 0.701 and a KS of 0.276, outperforming STAR’s 0.696 and 0.272, respectively, though the improvement is more modest. This suggests that while the GNN excels in graph-based relationship mining, its edge in binary classification may be influenced by the inherent complexity of default prediction.

To contextualize these results, the GBDT and GBDT_st baselines, rooted in traditional machine learning, lagged significantly, with AUCs of 0.986 and 0.987 for supply chain mining, and 0.574 and 0.581 for default prediction. The GAT and GAT_ensemble models, which incorporate attention mechanisms, showed incremental improvements (AUCs of 0.988 and 0.990 for supply chain, 0.671 and 0.673 for default prediction), yet they fell short of the GNN’s performance. The STAR model, designed for spatial-temporal tasks, provided a strong benchmark but was outpaced by our method, reinforcing the advantage of a tailored GNN architecture for static spatial modeling in this context.

7 Conclusion

We proposed a GNN-based framework for SME credit risk assessment by modeling supply chain interactions. Using large-scale real-world graphs (23.4M and 8.6M nodes), our method captured structural dependencies and outperformed baselines like GAT and STAR in supply chain mining and default prediction (AUCs: 0.995/0.701). Also as a computational engine for Financial Stability Oversight Council early-warning systems, this model can largely reduce crisis response latency in simulated scenarios of commodity financing

Table 3: Performance comparison of baselines

| | Supply chain | | Default Pred. | |
|--------------|--------------|-------|---------------|-------|
| | AUC | KS | AUC | KS |
| GBDT | | | | |
| GBDT_st | 0.986 | 0.917 | 0.574 | 0.094 |
| GAT | 0.987 | 0.921 | 0.581 | 0.102 |
| GAT_ensemble | 0.988 | 0.926 | 0.671 | 0.231 |
| STAR | 0.99 | 0.937 | 0.673 | 0.236 |
| GNN | 0.991 | 0.941 | 0.696 | 0.272 |

breakdowns. Banking regulators could deploy it to quantify contagion exposure thresholds under Basel III’s Pillar 2 requirements, directly strengthening the resilience of financial market utilities. While effective, limitations include static modeling, high costs, and noisy data. Future work may explore temporal GNNs, explainability (e.g., SHAP), and federated learning for enhanced accuracy and privacy.

References

- [1] Angilella, S., & Mazzù, S. (2019). A credit risk model with an automatic override for innovative small and medium-sized enterprises. *Journal of the Operational Research Society*, 70(10), 1784-1800.
- [2] Yoshino, N., Taghizadeh-Hesary, F., Charoensivakorn, P., & Niraula, B. (2016). Small and medium-sized enterprise (SME) credit risk analysis using bank lending data: An analysis of Thai SMEs. *Journal of Comparative Asian Development*, 15(3), 383-406.
- [3] Guo, F., Mo, H., Wu, J., Pan, L., Zhou, H., Zhang, Z., ... & Huang, F. (2024). A hybrid stacking model for enhanced short-term load forecasting. *Electronics*, 13(14), 2719.
- [4] Zhou, S., Zhang, Z., Zhang, R., Yin, Y., Chang, C. H., & Shen, Q. (2025, February). Regression and forecasting of us stock returns based on lstm. In *Proceedings of the 2025 3rd International Conference on Communication Networks and Machine Learning* (pp. 126-130).
- [5] Zhang, L., & Liang, R. (2025). Avocado price prediction using a hybrid deep learning model: TCN-MLP-attention architecture. *arXiv preprint arXiv:2505.09907*.
- [6] Wang, Y., Jia, P., Shu, Z., Liu, K., & Shariff, A. R. M. (2025). Multidimensional precipitation index prediction based on CNN-LSTM hybrid framework. *arXiv preprint arXiv:2504.20442*.
- [7] Wang, J., Zhang, Z., He, Y., Song, Y., Shi, T., Li, Y., ... & He, L. (2024). Enhancing Code LLMs with Reinforcement Learning in Code Generation. *arXiv preprint arXiv:2412.20367*.
- [8] Yang, H., Lyu, H., Zhang, T., Wang, D., & Zhao, Y. (2025). LLM-Driven E-Commerce Marketing Content Optimization: Balancing Creativity and Conversion. *arXiv preprint arXiv:2505.23809*.
- [9] Li, Z., He, S., Yang, Z., Ryu, M., Kim, K., & Madduri, R. (2024). Advances in applf: A comprehensive and extensible federated learning framework. *arXiv preprint arXiv:2409.11585*.
- [10] Zhong, J., & Wang, Y. (2025). Enhancing Thyroid Disease Prediction Using Machine Learning: A Comparative Study of Ensemble Models and Class Balancing Techniques.
- [11] Zheng, Z., Wu, S., & Ding, W. (2025). CTLformer: A Hybrid Denoising Model Combining Convolutional Layers and Self-Attention for Enhanced CT Image Reconstruction. *arXiv preprint arXiv:2505.12203*.
- [12] Zhong, J., Wang, Y., Zhu, D., & Wang, Z. (2025). A Narrative Review on Large AI Models in Lung Cancer Screening, Diagnosis, and Treatment Planning. *arXiv preprint arXiv:2506.07236*.
- [13] Lv, K., Wu, R., Chen, S., & Lan, P. (2024). CCI-YOLOv8n: Enhanced fire detection with CARAFE and context-guided modules. *arXiv preprint arXiv:2411.11011*.
- [14] Qiu, S., Li, Y., Wang, Z., Shen, Y., Li, Z., & Shen, F. (2024, September). Geometric Matrix Completion for Missing Data Estimation in Power Distribution Systems. In *2024 The 9th International Conference on Power and Renewable Energy (ICPRE)* (pp. 1605-1609). IEEE.
- [15] Li, Z. Retrieval-Augmented Forecasting with Tabular Time Series Data. In *The 4th Table Representation Learning Workshop at ACL 2025*.
- [16] Gu, J., Zhang, Y., & Hu, Z. (2025). FAN-TSF: A Frequency Adaptive Normalization Approach for Non-stationary Time Series Forecasting on Stock Market Data. Available at SSRN 5259125.

- [17] He, Y., Li, S., Wang, J., Li, K., Song, X., Yuan, X., ... & Wang, X. (2025). Enhancing low-cost video editing with lightweight adaptors and temporal-aware inversion. arXiv preprint arXiv:2501.04606.
- [18] Zeng, S., Chang, X., Xie, M., Liu, X., Bai, Y., Pan, Z., ... & Wei, X. (2025). FutureSightDrive: Thinking Visually with Spatio-Temporal CoT for Autonomous Driving. arXiv preprint arXiv:2505.17685.
- [19] Chen, K., Qiu, L., Xie, X., & Bai, Y. (2025). Dynasmile: Video-based smile analysis software in orthodontics. *SoftwareX*, 29, 102004.
- [20] Zhou, Z. (2025). Research on the Application of Intelligent Robots and Software in Multiple Segmentation Scenarios Based on Machine Learning. Available at SSRN 5236930.
- [21] Wu, S., Fu, L., Chang, R., Wei, Y., Zhang, Y., Wang, Z., ... & Li, K. (2025). Warehouse robot task scheduling based on reinforcement learning to maximize operational efficiency. *Authorea Preprints*.
- [22] Yu, D., Liu, L., Wu, S., Li, K., Wang, C., Xie, J., ... & Ji, R. (2025, March). Machine learning optimizes the efficiency of picking and packing in automated warehouse robot systems. In 2025 IEEE International Conference on Electronics, Energy Systems and Power Engineering (EESPE) (pp. 1325-1332). IEEE.
- [23] Zhao, Y., Peng, Y., Li, D., Yang, Y., Zhou, C., & Dong, J. (2025). Research on Personalized Financial Product Recommendation by Integrating Large Language Models and Graph Neural Networks. arXiv preprint arXiv:2506.05873.
- [24] Jin, Y., & Yang, Z. (2025, January). Scalability optimization in cloud-based ai inference services: Strategies for real-time load balancing and automated scaling. In *Proceedings of the 2025 4th International Conference on Big Data, Information and Computer Network* (pp. 266-270).
- [25] Wu, D., Guo, F., Yao, Z., Zhu, D., Zhang, Z., Li, L., ... & Zhang, J. (2024). Enhancing Reliability and Performance of Load Frequency Control in Aging Multi-Area Power Systems under Cyber-Attacks. *Applied Sciences*, 14(19), 8631.
- [26] Chen, Y., Liu, L., & Fang, L. (2024). An Enhanced Credit Risk Evaluation by Incorporating Related Party Transaction in Blockchain Firms of China. *Mathematics*, 12(17), 2673.
- [27] Zhang, Z., Li, X., Cheng, Y., Chen, Z., & Liu, Q. (2025). Credit Risk Identification in Supply Chains Using Generative Adversarial Networks. arXiv preprint arXiv:2501.10348.
- [28] Hu, Z., Yu, R., Zhang, Z., Zheng, H., Liu, Q., & Zhou, Y. (2024). Developing Cryptocurrency Trading Strategy Based on Autoencoder-CNN-GANs Algorithms. arXiv preprint arXiv:2412.18202.
- [29] Ouyang, K., Fu, S. (2024). Graph Neural Networks Are Evolutionary Algorithms. arXiv preprint arXiv:2412.17629.
- [30] Cao, Z., Zhao, S., Dong, Z., & Douady, R. (2024). Financial Network Analysis Using Polymodel Theory.
- [31] Yu, Q., Yin, Y., Zhou, S., Mu, H., & Hu, Z. (2025). Detecting Financial Fraud in Listed Companies via a CNN-Transformer Framework.
- [32] Li, Z. Episodic Memory Banks for Lifelong Robot Learning: A Case Study Focusing on Household Navigation and Manipulation. In *Workshop on Foundation Models Meet Embodied Agents at CVPR 2025*.
- [33] Wu, C., Huang, H., Chen, J., Zhou, M., & Han, S. (2024). A novel Tree-augmented Bayesian network for predicting rock weathering degree using incomplete dataset. *International Journal of Rock Mechanics and Mining Sciences*, 183, 105933.
- [34] Qiu, S., Wang, Y., Ke, Z., Shen, Q., Li, Z., Zhang, R., & Ouyang, K. (2025). A generative adversarial network-based investor sentiment indicator: Superior predictability for the stock market. *Mathematics*, 13(9), 1476.
- [35] Zhao, S., Wang, D., & Douady, R. (2024). PolyModel for Hedge Funds' Portfolio Construction Using Machine Learning. arXiv preprint arXiv:2412.11019.
- [36] Wang, H., Gong, Y., & Yu, C. (2025). GAN_BERT: An Advanced Neural Architecture for Effective Fraud Detection on Imbalanced Datasets.
- [37] Bian, W. (2024). An Optimal Control Approach for Inverse Problems with Deep Learnable Regularizers. arXiv preprint arXiv:2409.00498.
- [38] Yang, H., Lu, Q., Wang, Y., Liu, S., Zheng, J., & Xiang, A. (2025). User Behavior Analysis in Privacy Protection with Large Language Models: A Study on Privacy Preferences with Limited Data. arXiv preprint arXiv:2505.06305.
- [39] Luo, J., Song, Q., Guo, F., Wu, H., Som, H. M., Alahmari, S., & Hoshyar, A. N. (2025). Joint deep reinforcement learning strategy in MEC for smart internet of vehicles edge computing networks. *Sustainable Computing: Informatics and Systems*, 46, 101121.
- [40] Lu, Q., Lyu, H., Zheng, J., Wang, Y., Zhang, L., & Zhou, C. (2025). Research on E-Commerce Long-Tail Product Recommendation Mechanism Based on Large-Scale Language Models. arXiv preprint arXiv:2506.06336.
- [41] Zhang, Y., Mei, D., Luo, H., Xu, C., & Tsai, R. T. H. (2025). Smutf: Schema matching using generative tags and hybrid features. *Information Systems*, 102570.
- [42] Zhang, Yuxin and Chen, Naxi and Zhang, Yeyubei and Wu, Wenjun, Research on Business Decision Support System based on Big Data and Artificial Intelligence (June 30, 2025). Available at SSRN: <http://dx.doi.org/10.2139/ssrn.5332298>
- [43] Li, Y., Wu, D., Chen, J., Shi, W., Wang, L., Su, L., ... & Zhang, D. (2024). SigCan: Toward Reliable ToF Estimation Leveraging Multipath Signal Cancellation on Commodity WiFi Devices. *IEEE Transactions on Mobile Computing*.
- [44] Shi, W., Le, H. V., & Choo, K. T. W. (2025, April). Towards Multimodal Large-Language Models for Parent-Child Interaction: A Focus on Joint Attention. In *Extended Abstracts of the 2025 CHI Conference on Human Factors in Computing Systems (CHI EA '25)* (pp. 1–6). Association for Computing Machinery. <https://doi.org/10.1145/3706599.3720215>
- [45] Li, Q., Luo, T., Jiang, M., Jiang, Z., Hou, C., & Li, F. (2025, April). Semi-Supervised Multi-View Multi-Label Learning with View-Specific Transformer and Enhanced Pseudo-Label. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 39, No. 17, pp. 18430-18438).
- [46] Li, Q., Luo, T., Jiang, M., Liao, J., & Jiang, Z. (2024, October). Deep Incomplete Multi-View Network Semi-Supervised Multi-Label Learning with Unbiased Loss. In *Proceedings of the 32nd ACM International Conference on Multimedia* (pp. 9048-9056).