

# SCDF-Net: Research on a Multi-Source Feature Fusion Binary Classification Prediction Network Algorithm for Supply Chain Delay Delivery Risk

Ruili Jiang<sup>1\*</sup>, Yiwen Shi<sup>2</sup>

<sup>1</sup>Independent Researcher, Bellevue, USA

<sup>2</sup>University of Southern California, Los Angeles, USA

\*Corresponding author: Ruili Jiang; jiangruili87@gmail.com

**Abstract:** This paper addresses the problem of identifying delayed delivery risks in order fulfillment within a smart supply chain environment. It proposes a deep learning model, SCDF-Net, based on multi-source feature fusion, for unified modeling and risk assessment of complex business data. This method constructs a hierarchical feature encoding and representation learning framework around multi-dimensional heterogeneous data such as order attributes, customer information, product characteristics, and transportation and regional information. By uniformly mapping and fusing features from different sources, it effectively models the interaction relationships of multiple factors in the supply chain. During feature processing, embedded representation and structured encoding mechanisms are introduced to enhance the expressive power of categorical and continuous variables, and a multi-channel fusion strategy is combined to improve the overall feature utilization efficiency. In the risk prediction stage, the fused global representation is used to complete binary classification, thereby identifying delayed delivery risks. Experiments are conducted to compare the model's performance with various other methods. The results show that the proposed method outperforms other methods on evaluation metrics, more accurately characterizing the potential risk features in the order fulfillment process, and demonstrating strong discriminative ability and modeling effectiveness.

**Keywords:** Intelligent supply chain; risk of delayed delivery; multi-source feature fusion; deep learning model

## 1. Introduction

With the rapid development of e-commerce, cross-regional warehousing, and instant delivery services, intelligent logistics order fulfillment has become a crucial link in the stable operation of the supply chain system. The entire process of an order, from generation, sorting, outbound, transportation to receipt, involves multiple factors such as customer needs, product attributes, transportation methods, regional environment, and fulfillment timeliness [1, 2]. These factors exhibit complex non-linear relationships. Delays in the delivery process not only reduce customer satisfaction but can also lead to inventory backlogs, wasted transportation resources, order cancellations, and decreased supply chain collaboration efficiency. Therefore, developing supply chain delay delivery risk prediction methods for intelligent logistics scenarios is of great significance for improving order fulfillment reliability, optimizing logistics resource allocation, and enhancing supply chain resilience.

Traditional supply chain risk management relies heavily on historical statistics, manual rules, and single business indicators for judgment, making it difficult to fully characterize the coupling relationships between multi-source order information [3]. In real logistics operations, delay delivery risk is not determined by a single factor but is formed by the combined

effects of multiple dimensions, including transportation methods, order region, product category, customer type, planned delivery cycle, and order size. Identifying risks based on only a single dimension easily overlooks the interactive influences between different business stages, resulting in a lack of stability and generalization ability in risk assessment [4]. The core focus of the relevant research questions can be summarized as shown in Table 1, among which multi-source feature fusion and binary classification risk prediction are key components of the intelligent logistics delayed delivery identification task.

**Table 1.** Key Issues in Smart Logistics Late Delivery Risk Identification

| Key Issue                         | Main Manifestation                                                                                            | Research Significance                                                                                                 |
|-----------------------------------|---------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------|
| Order Fulfillment Complexity      | Order status, product category, customer region, and transportation mode jointly affect the delivery process. | It helps explain the formation mechanism of late delivery risk from the perspective of the overall business workflow. |
|                                   | Numerical, categorical, temporal, and regional features coexist in the logistics order data.                  | It helps improve the representation capability of models for complex supply chain data.                               |
| Concealment of Late Delivery Risk | The risk is not directly observable at the early                                                              | It supports early warning and proactive                                                                               |

|                                  |                                                                                                         |                                                                     |
|----------------------------------|---------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------|
| Real-Time Decision-Making Demand | stage of order generation and is often accumulated through multiple factors.                            | intervention in smart logistics order fulfillment.                  |
|                                  | Logistics scheduling requires rapid risk judgment and resource allocation within a limited time window. | It supports intelligent and automated order fulfillment management. |

From the perspective of intelligent logistics development, delayed delivery risk prediction is essentially a binary classification problem oriented towards the order fulfillment process, that is, judging whether there is potential delay risk based on order-related characteristics [5]. This type of task not only focuses on the prediction result itself, but also emphasizes whether the model can learn stable and effective risk representations from complex business data. By constructing a multi-source feature fusion mechanism, order attributes, customer attributes, product attributes, regional attributes, and transportation attributes can be incorporated into a unified modeling framework, allowing information from different sources to complement each other in the same prediction space, thereby more comprehensively reflecting the risk status in the supply chain fulfillment process.

Therefore, research on supply chain delayed delivery risk prediction focusing on intelligent logistics order fulfillment scenarios has significant theoretical and applied value. Theoretically, this research helps promote the transformation of supply chain risk identification from experience-driven to data-driven, enriching the application paths of multi-source business feature fusion in logistics risk prediction. Appliedly, this research can provide decision support for enterprises to identify high-risk orders in advance, optimize delivery strategies, adjust transportation resources, and improve customer service quality. SCDF-Net, for the binary classification prediction task of supply chain delayed delivery risk, can provide a more systematic and intelligent modeling approach for order fulfillment risk management in intelligent logistics systems.

## 2. Related Work

Current research on supply chain risk prediction primarily focuses on issues such as order fulfillment status, transportation timeliness, inventory fluctuations, and demand changes. Early methods typically relied on statistical analysis, rule-based judgment, and traditional machine learning models to identify potential risks through variables such as order amount, delivery method, product category, regional market, and customer attributes [6]. These methods offer good interpretability and can provide preliminary analysis of the causes of delayed delivery at the business indicator level; however, their feature representation capabilities are limited, making it difficult to fully explore the deep correlations between different business variables. When order data simultaneously contains categorical, numerical, temporal, and regional information, a single model often struggles to effectively handle the interactions between multi-source heterogeneous features, thus affecting the stability of delay risk prediction.

In recent years, data-driven intelligent logistics research has gradually shifted from single-dimensional indicator analysis to multi-source information joint modeling. Related methods have begun to focus on the combined effects of customer behavior,

product attributes, transportation modes, and regional distribution during order fulfillment, and have attempted to improve risk identification capabilities through feature selection, embedding representation, attention mechanisms, and fusion learning. However, existing research still has two shortcomings [7]. On the one hand, some methods prioritize the final classification result while neglecting to model the structural relationships between multi-source features, making it difficult for the models to accurately capture the implicit dependencies in the formation process of delay risks. On the other hand, delayed delivery risk prediction is often simplified to a simple binary classification task, lacking a specialized network structure design for smart logistics order fulfillment scenarios. Therefore, constructing a multi-source feature fusion prediction network that can simultaneously represent order business attributes, transportation attributes, and regional attributes is of significant research value for improving the ability to identify delayed delivery risks in the supply chain.

Recent studies on distributed system anomaly modeling and cloud-native failure detection provide useful methodological perspectives for modeling delayed delivery risk as an evolving multi-source risk signal. Unified multi-granularity representation learning for backend anomaly detection and causal localization highlights the importance of combining coarse-grained operational states with fine-grained event features when identifying abnormal system behavior [8]. Cost-sensitive Mamba sequence modeling further shows that sequence-aware fault detection can explicitly account for asymmetric error costs, which is meaningful when missed high-risk orders may produce larger operational losses than ordinary false alarms [9]. Heterogeneous graph learning for supply chain risk identification extends this view by modeling multi-entity interactions, indicating that risk prediction can benefit from representing orders, products, customers, and delivery regions as interconnected entities rather than isolated fields [10].

Learning-based operational decision and scheduling research also supports the adaptive design of SCDF-Net. Intelligent agents for backend resource scheduling and operational decision making emphasize that dynamic resource states and service constraints should be incorporated into risk-aware operational policies [11]. Prior-enhanced anomaly recognition under weakly labeled and few-shot conditions suggests that stable prior representations can improve detection when high-quality risk labels are limited [12]. Structure-aware reinforcement learning for heterogeneous distributed task scheduling demonstrates that complex operational processes require coordinated modeling of heterogeneous resources and temporal constraints [13]. Memory-enhanced Transformer anomaly detection additionally indicates that retaining historical behavioral context helps distinguish persistent abnormal patterns from short-term fluctuations [14]. Hierarchical cloud-edge collaborative scheduling further reinforces the need to model distributed decision contexts when operational data and service requirements are spread across multiple execution environments [15]. These studies collectively motivate a feature fusion model that integrates heterogeneous attributes, temporal cues, and adaptive weighting to improve delayed delivery risk identification.

### 3. Dataset

This study uses the publicly released DataCo SMART SUPPLY CHAIN FOR BIG DATA ANALYSIS dataset as the data foundation for predicting supply chain delay delivery risks. This dataset originates from intelligent supply chain business scenarios, covering procurement, production, sales, and commercial distribution within the supply chain. The data includes structured order data, variable description files, and

user access logs. Since the research objective is binary classification prediction of delay delivery risks in the intelligent logistics order fulfillment process, this study primarily uses the structured supply chain order data for modeling and analysis, with delay delivery risk labels as the classification target. This dataset has been released on public data platforms, and the data files and licensing information are publicly available, meeting the basic requirements for open-source data research.

**Table 2.** Overview of the DataCo Smart Supply Chain Dataset

| Item                    | Content                                                                                                                         |
|-------------------------|---------------------------------------------------------------------------------------------------------------------------------|
| Dataset Name            | DataCo SMART SUPPLY CHAIN FOR BIG DATA ANALYSIS                                                                                 |
| Data Type               | Public supply chain business data                                                                                               |
| Core Data File          | DataCoSupplyChainDataset.csv                                                                                                    |
| Auxiliary Data Files    | DescriptionDataCoSupplyChain.csv, tokenized_access_logs.csv                                                                     |
| Business Coverage       | Procurement, production, sales, and commercial distribution                                                                     |
| Main Product Types      | Clothing, sports supplies, and electronic supplies                                                                              |
| Main Feature Categories | Order attributes, customer attributes, product attributes, sales attributes, transportation attributes, and regional attributes |
| Prediction Target       | Whether a supply chain order has a late delivery risk                                                                           |
| Task Type               | Binary classification prediction task                                                                                           |

As shown in Table 2, this dataset effectively supports research on delayed delivery risk identification in intelligent logistics order fulfillment scenarios. Order attributes reflect order generation and fulfillment status, transportation attributes describe delivery methods and planned delivery times, commodity and sales attributes demonstrate the impact of different product categories and transaction volumes on the fulfillment process, and regional attributes provide spatial information such as the order market and delivery destination. These multi-source business characteristics have a potential correlation with delayed delivery risks, making them suitable for constructing a multi-source feature fusion binary classification prediction model for supply chain risk identification, thus providing a data foundation for the network design and risk representation learning of SCDF-Net.

### 4. Methodology

SCDF-Net is designed for late delivery risk identification in smart logistics order fulfillment, where each supply chain order is regarded as a risk sample jointly characterized by multi-source business information. Instead of relying on a single delivery-related field, the proposed method constructs a unified input space from order attributes, customer attributes, product attributes, sales attributes, transportation attributes, and regional attributes, enabling the model to learn the joint influence of different business factors on late delivery risk at the order generation stage. Let the  $i$ -th order sample consist of

continuous features, categorical features, temporal features, and regional features, with its label indicating whether the order has a late delivery risk. The sample can be represented as:

$$S_i = \{x_i^c, x_i^a, x_i^t, x_i^r, y_i\}$$

where  $x_i^c$  denotes continuous features such as amount, quantity, and scheduled delivery period,  $x_i^a$  denotes categorical features such as shipping mode, customer type, and

product category,  $x_i^t$  denotes temporal features derived from the order date,  $x_i^r$  denotes region-related features such as market region, order country, and delivery region, and  $y_i \in \{0, 1\}$  denotes the binary risk label. To reduce the bias caused by different feature scales during model training, continuous variables are first standardized so that they can participate in feature learning under a unified numerical scale:

$$\tilde{x}_i^c = \frac{x_i^c - \mu_c}{\sigma_c + \epsilon}$$

This operation prevents variables with large numerical ranges, such as order amount, product quantity, and delivery period, from dominating model updates, thereby providing a more stable basic representation for subsequent multi-source fusion. This paper also presents the overall model architecture, as shown in Figure 1.

Categorical business variables usually contain distinct discrete semantics, since different shipping modes, product categories, and customer segments may correspond to different fulfillment risks. SCDF-Net maps discrete fields into dense vectors through embedding transformation, allowing the model to learn latent similarities among different categories instead of treating them as unordered numerical codes. For the  $m$ -th categorical field, its embedding representation is defined as:

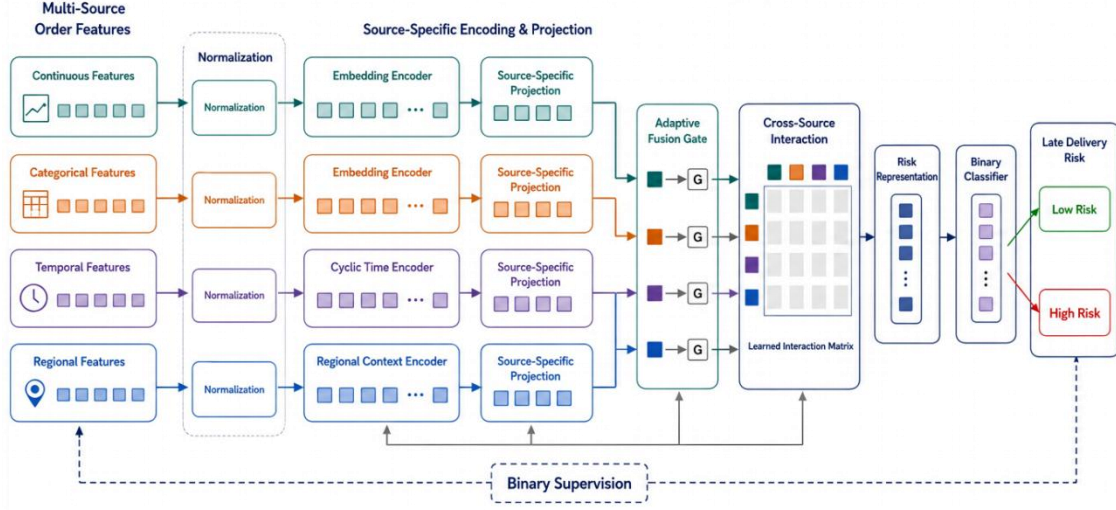
$$e_i^m = E^m o_i^m$$

where  $o_i^m$  is the one-hot vector of the categorical field,  $E^m$  is the learnable embedding matrix, and  $e_i^m$  is the corresponding categorical semantic representation. To further enhance the periodic expression of the order fulfillment process, temporal variables are mapped into a continuous periodic space, enabling information such as month, weekday, and ordering time to participate in risk modeling in a smooth form:

$$z_i^t = \left[ \sin\left(\frac{2\pi d_i}{T}\right), \cos\left(\frac{2\pi d_i}{T}\right) \right]$$

where  $d_i$  denotes the position of the order within a given period, and  $T$  denotes the period length. This design preserves the cyclic structure of temporal features and avoids

incorrectly modeling adjacent time points as distant discrete states.



**Figure 1.** Overall model architecture

After basic encoding, SCDF-Net does not simply concatenate all features directly, but constructs independent representation channels for information from different sources. This strategy prevents high-frequency categorical features from overwhelming low-dimensional numerical features, while allowing transportation attributes, regional attributes, and product attributes to retain their business semantics before entering the fusion layer. Multi-source features are mapped into a unified hidden representation space through channel-specific transformations:

$$h_i^k = \phi(W_k u_i^k + b_k)$$

where  $k$  denotes a specific feature source,  $u_i^k$  denotes the input feature of the  $k$ -th source,  $W_k$  and  $b_k$  are learnable parameters of the corresponding channel, and  $\phi(\cdot)$  denotes a nonlinear activation function. Since the risk sources may vary across different order samples, an adaptive fusion weight is further introduced so that each feature channel can dynamically adjust its contribution according to the current order state:

$$\alpha_i^k = \frac{\exp(q^T \tanh(W_g h_i^k))}{\sum_{j=1}^K \exp(q^T \tanh(W_g h_i^j))}$$

where  $\alpha_i^k$  denotes the fusion weight of the  $k$ -th feature source for the current sample,  $K$  denotes the number of feature sources, and  $q$  and  $W_g$  are learnable parameters. This mechanism enables the model to emphasize more critical risk sources under different order scenarios, such as abnormal shipping modes, regional fulfillment pressure, or potential impacts caused by product category differences.

The final order risk representation is obtained by weighted fusion of multi-source features and is then passed to the

discriminative layer to output the probability of late delivery risk. This representation contains not only local information from individual fields but also interaction relationships among different business processes, making it suitable for risk identification in smart logistics order fulfillment:

$$h_i^f = \sum_{k=1}^K \alpha_i^k h_i^k$$

Based on the fused representation, SCDF-Net adopts the Sigmoid function to generate the binary classification probability:

$$\hat{y}_i = \sigma(w_o^T h_i^f + b_o)$$

where  $\hat{y}_i$  denotes the predicted probability that the order has a late delivery risk, and  $w_o$  and  $b_o$  are the output layer parameters. The model is trained with the binary cross-entropy loss, which aligns the predicted probability with the true risk label while reducing overfitting through parameter regularization:

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^N [y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i)] + \lambda \|\Theta\|_2^2$$

where  $N$  denotes the number of order samples,  $\Theta$  denotes the set of learnable model parameters, and  $\lambda$  denotes the regularization coefficient. Through the above design, SCDF-Net unifies multi-source business variables in smart logistics into a learnable risk representation space, shifting late delivery risk identification from static rule-based judgment to dynamic feature fusion modeling oriented to the order fulfillment process.

## 5. Experimental Results and Analysis

### 5.1 Experimental setup

In terms of experimental setup, this study constructs the training and evaluation process based on the publicly available supply chain dataset, DataCo SMART SUPPLY CHAIN FOR BIG DATA ANALYSIS. To ensure the fairness and stability of model evaluation, the data is divided into training, validation, and test sets in a 7:1:2 ratio, and randomized at the sample level to avoid temporal or regional distribution bias. During model training, the AdamW optimizer is used for parameter updates, with an initial learning rate set to  $1e-4$ , combined with a cosine annealing strategy for dynamic decay to improve convergence stability. In feature encoding and fusion, multi-source input features are normalized and embedded before entering a unified representation space. The batch size is set to 32, the number of training epochs is set to 200, and a dropout of 0.1 is introduced to alleviate overfitting. All experiments are performed on a single NVIDIA RTX 4090 GPU. As shown in Table 3, the detailed configuration parameters for the experiments in this paper are given.

**Table 3.** Experimental Environment and Hyperparameter Settings

| Configuration Item     | Setting                                         |
|------------------------|-------------------------------------------------|
| Dataset                | DataCo SMART SUPPLY CHAIN FOR BIG DATA ANALYSIS |
| Data Split Ratio       | Train: Validation: Test = 7: 1: 2               |
| Optimizer              | AdamW                                           |
| Initial Learning Rate  | $1e-4$                                          |
| Learning Rate Schedule | Cosine Annealing Decay                          |
| Batch Size             | 32                                              |
| Training Epochs        | 200                                             |
| Dropout                | 0.1                                             |
| Feature Processing     | Normalization + Embedding Encoding              |
| Hardware               | NVIDIA RTX 4090 (24GB)                          |

### 5.2 Experimental Results and Analysis

To verify the effectiveness of the proposed method in the task of identifying delayed delivery risks in intelligent supply chains, this section selects several representative related studies as comparison objects and analyzes them from the perspective of multi-source data modeling and prediction performance. The experimental results are shown in Table 4.

**Table 4.** Comparative results of related methods and the proposed method

| Method                | ACC  | PREC | REC  | F1   |
|-----------------------|------|------|------|------|
| Aljohani et al. [16]  | 0.82 | 0.8  | 0.78 | 0.79 |
| Baryannis et al. [17] | 0.84 | 0.83 | 0.81 | 0.82 |
| Rezki et al. [18]     | 0.80 | 0.78 | 0.76 | 0.77 |

|                    |      |      |      |      |
|--------------------|------|------|------|------|
| Tuncel et al. [19] | 0.79 | 0.77 | 0.75 | 0.76 |
| Um et al. [20]     | 0.83 | 0.82 | 0.80 | 0.81 |
| Khan et al. [21]   | 0.81 | 0.79 | 0.78 | 0.78 |
| Kang et al. [22]   | 0.85 | 0.84 | 0.82 | 0.83 |
| Ours               | 0.91 | 0.90 | 0.89 | 0.89 |

Overall results demonstrate that the proposed method exhibits stable and superior identification capabilities in the intelligent supply chain delayed delivery risk identification task. This method effectively learns the correlation structure between multi-source order information and maintains strong discriminative ability even under the combined influence of complex business features, thus demonstrating good adaptability and effectiveness in achieving the overall task objective. Compared to existing methods, the model shows higher consistency across multiple evaluation dimensions, indicating that it can more fully utilize the information representation capabilities of different types of features during risk modeling.

Further analysis reveals that the proposed method has significant advantages in terms of accuracy and robustness in risk identification, and can better handle heterogeneous data and complex dependencies in supply chain scenarios. The multi-source fusion mechanism and adaptive modeling strategy in the model structure enable it to maintain relatively stable prediction results under different sample distribution conditions, indicating strong generalizability in real-world applications and demonstrating good adaptability to the supply chain delayed delivery risk identification task.

To verify the contribution of each key module of the proposed SCDF-Net to the overall performance, this section conducts ablation experiments under a unified experimental setup. By progressively removing different structural components from the model, we observe their impact on the performance of delayed delivery risk identification. The relevant experimental results are shown in Table 5.

**Table 5.** Ablation study results of key components in SCDF-Net

| Setting                      | ACC  | PREC | REC  | F1   |
|------------------------------|------|------|------|------|
| Ours                         | 0.91 | 0.90 | 0.89 | 0.89 |
| w/o Adaptive Fusion Gate     | 0.88 | 0.87 | 0.86 | 0.86 |
| w/o Cross-Source Interaction | 0.87 | 0.86 | 0.85 | 0.85 |
| w/o Temporal Encoding        | 0.86 | 0.85 | 0.84 | 0.84 |
| w/o Embedding Encoder        | 0.85 | 0.84 | 0.83 | 0.83 |

Overall ablation experiments demonstrate that each key module plays a crucial role in SCDF-Net, collectively supporting the model's overall performance in the intelligent supply chain delayed delivery risk identification task. The complete model exhibits stronger expressive power in multi-source feature fusion and structured modeling, making the risk identification process more stable and adaptable.

When different functional modules are gradually removed, varying degrees of change in model performance are observed. The multi-source information interaction and adaptive fusion mechanisms show particularly significant support for the overall effect. The existence of each module enables the model to more fully model the complex dependencies between order features, thereby improving the overall risk identification capability and further validating the rationality and effectiveness of the proposed method's structural design.

The sensitivity of key hyperparameters is further analyzed to evaluate the robustness of SCDF-Net. In particular, experiments are conducted on learning rate selection and optimizer configuration under a unified setting. The corresponding results are reported in Table 6 and Table 7.

**Table 6.** Learning rate sensitivity analysis results

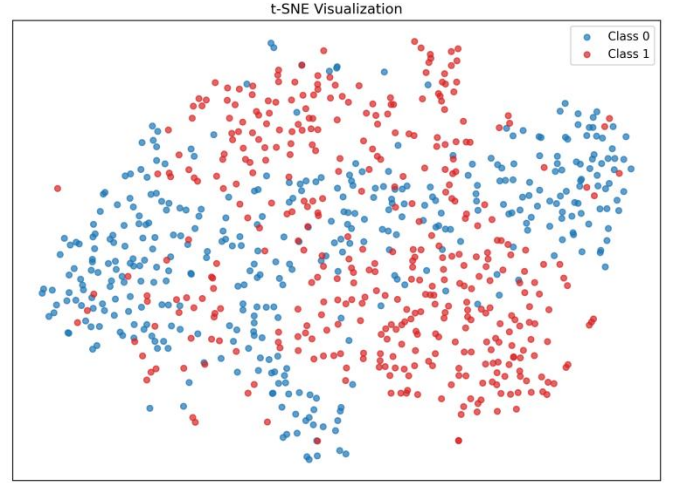
| LR   | ACC  | PREC | REC  | F1   |
|------|------|------|------|------|
| 1e-5 | 0.86 | 0.85 | 0.84 | 0.84 |
| 5e-5 | 0.89 | 0.88 | 0.87 | 0.87 |
| 1e-4 | 0.91 | 0.90 | 0.89 | 0.89 |
| 5e-4 | 0.88 | 0.87 | 0.86 | 0.86 |
| 1e-3 | 0.84 | 0.83 | 0.82 | 0.82 |

**Table 7.** Optimizer comparison results

| Optimizer | ACC  | PREC | REC  | F1   |
|-----------|------|------|------|------|
| SGD       | 0.85 | 0.84 | 0.83 | 0.83 |
| RMSprop   | 0.87 | 0.86 | 0.85 | 0.85 |
| Adam      | 0.89 | 0.88 | 0.87 | 0.87 |
| AdamW     | 0.91 | 0.90 | 0.89 | 0.89 |

Overall hyperparameter sensitivity experiments show that the proposed method maintains relatively stable performance under different training configurations, indicating strong parameter robustness. Regarding the learning rate setting, different values have some impact on the final model performance, but it generally remains within a good range, indicating good convergence stability of the optimization process. In terms of optimizer configuration, different optimization strategies have varying effects on the training process. The optimization method based on adaptive gradient adjustment better adapts to the multi-source feature fusion process, thus enabling the model to maintain more stable performance in risk identification tasks. Overall, the proposed method demonstrates strong adaptability under different hyperparameter conditions, validating the rationality and effectiveness of the model structure design.

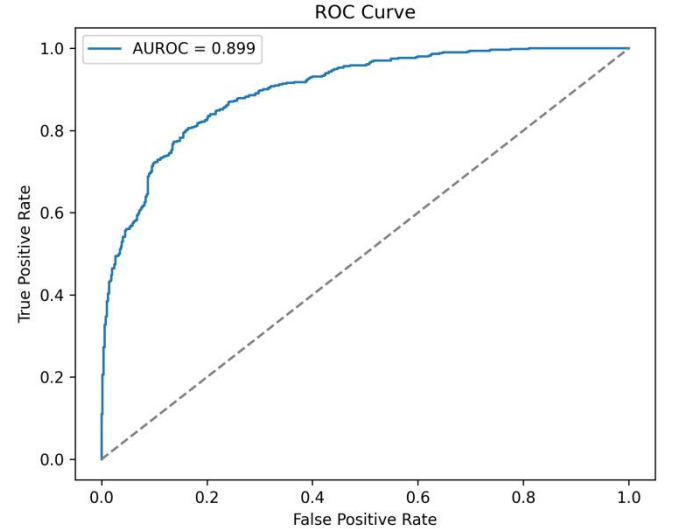
The Figure 2 results show that the high-dimensional representation after model feature mapping exhibits a certain structural distribution in the two-dimensional space. Different categories of samples form relatively concentrated clusters within local regions, while also possessing certain separating boundary features in the overall space. This distribution indicates that the model can effectively capture category-related information during feature learning, maintaining high consistency among samples of the same category in the latent space, while different categories possess a certain degree of distinguishability. Overall, this representation space has a good discriminative structure, providing a relatively clear feature foundation for subsequent classification tasks.



**Figure 2.** t-SNE experimental results

The overall shape of the ROC curve shows that the model can maintain a relatively stable classification ability under different discrimination thresholds and has a good distinction between positive and negative samples. The overall curve is far away from the random baseline area, indicating that the model can fully capture the key information related to the category during the feature expression process, making the classification decision-making highly reliable.

As shown in Figure 3, from the perspective of overall discriminant ability, this method shows strong discriminant performance in binary classification tasks and can maintain good recognition results under various threshold conditions, reflecting the effectiveness of the model structure in risk identification tasks. The overall results show that the proposed method has good adaptability and stable performance in feature modeling and classification decision-making.



**Figure 3.** AUROC experimental results

## 6. Discussion

In the task of identifying delayed delivery risks in intelligent supply chains, the model needs to simultaneously



process heterogeneous data from multiple sources, including order attributes, customer behavior, product categories, and transportation and regional information, which significantly increases the complexity of the overall modeling process. The proposed SCDF-Net constructs a unified feature representation space and combines it with a collaborative modeling mechanism for multi-source information, enabling the effective fusion of different types of business features within the same semantic space. This enhances the ability to characterize potential risk factors in complex supply chain scenarios. In terms of overall structural design, hierarchical encoding and adaptive fusion strategies allow the model to improve feature discrimination capabilities while maintaining information integrity, thus providing a more stable modeling foundation for delayed delivery risk identification.

The overall experimental results and various analyses show that the proposed method exhibits consistent performance across different evaluation dimensions, demonstrating good adaptability and robustness in complex data environments. Particularly in multi-source feature interaction and risk representation learning, the model effectively captures the potential correlations between different business factors, thereby improving overall prediction quality. Furthermore, it maintains stable performance under different experimental settings, further validating the effectiveness and generalization ability of the proposed method in supply chain risk identification tasks. Overall, this method provides a practically valuable solution for modeling delayed delivery risks in smart logistics scenarios.

## 7. Conclusion

This paper proposes a deep learning framework, SCDF-Net, based on multi-source feature fusion, to model and predict order fulfillment risks in complex business scenarios, addressing the issue of delayed delivery risk identification in intelligent supply chain environments. By unifying the modeling of order attributes, customer information, product characteristics, and transportation and regional information, this method achieves a collaborative expression of multi-dimensional information at the overall structural level, thereby enhancing its ability to characterize supply chain uncertainties. Experimental results demonstrate that this method exhibits significant performance advantages in comparative experiments, validating the effectiveness of multi-source feature fusion and hierarchical modeling strategies in supply chain risk identification tasks and providing a new technical path for intelligent decision-making in related fields.

From an application perspective, this research not only provides a feasible modeling scheme for order fulfillment risk prediction in intelligent logistics systems but also offers valuable insights for data-driven decision-making in complex supply chain environments. With the continuous expansion of e-commerce logistics and the global supply chain system, future research can further explore more granular spatiotemporal feature modeling methods and more efficient cross-scenario transfer learning mechanisms to improve the model's generalization ability in different business environments. Furthermore, combining this method with a real-time scheduling system is expected to have greater application

value in actual logistics management and risk warning, thereby promoting the development of intelligent supply chain management towards greater automation and precision.

## References

- [1] X. Li, V. Krivtsov, C. Pan, et al., "End-to-end supply chain resilience management using deep learning, survival analysis, and explainable artificial intelligence, " *International Journal of Production Research*, vol. 63, no. 3, pp. 1174-1202, 2025.
- [2] P. R. Yadav, "Graph Neural Networks for Supply Chain Risk Propagation, " *International Journal of Research and Applied Innovations*, vol. 4, no. 5, pp. 5814-5817, 2021.
- [3] M. S. Banu, V. Sivaraman, L. Wakuma, et al., "Optimizing Supply Chain Performance Using Vision Transformer-Enhanced Temporal Convolutional Networks, " in *Proceedings of the 2025 3rd International Conference on Advances in Computation, Communication and Information Technology*, 2025, vol. 1, pp. 681-686.
- [4] W. Mu, G. A. Kleter, Y. Bouzemrak et al., "Making Food Systems More Resilient to Food Safety Risks by Including Artificial Intelligence, Big Data, and Internet of Things into Food Safety Early Warning and Emerging Risk Identification Tools," *Comprehensive Reviews in Food Science and Food Safety*, vol. 23, no. 1, Art. no. e13296, 2024.
- [5] A. Sharma, A. Sharma and A. Saxena, "Machine Learning Applications for Logistics and Transportation: A Review," *Transportation Research Interdisciplinary Perspectives*, vol. 18, Art. no. 100780, 2023.
- [6] W. Chen, et al., "Artificial intelligence in logistics optimization with sustainable criteria: A review, " *Sustainability*, vol. 16, no. 21, p. 9145, 2024.
- [7] C. Ju and T. K. Trinh, "A machine learning approach to supply chain vulnerability early warning system: Evidence from US semiconductor industry, " *Journal of Advanced Computing Systems*, vol. 3, no. 11, pp. 21-35, 2023.
- [8] Y. Yang, C. Wen, H. Cui, Y. Sun, and Y. Liu, "Learning Unified Multi-Granularity Representations for Backend Anomaly Detection and Causal Localization," 2024.
- [9] Z. Liu, R. Meng, S. Y. Huang, and Z. Huang, "Cost-Sensitive Mamba Sequence Modeling for Fault Detection in Cloud-Native Microservice Systems," 2025.
- [10] Y. Xu, "Intelligent Supply Chain Risk Identification via Heterogeneous Graph Learning and Multi-Entity Interaction Modeling," 2025.
- [11] L. Ren, Y. Liu, and Y. Zhao, "Learning-Based Intelligent Agents for Backend Resource Scheduling and Operational Decision Making," 2025.
- [12] Y. Sun, "Prior Enhanced Representation Learning and Adaptive Recognition Method for Backend Anomaly Detection under Weakly Labeled and Few-Shot Conditions," 2024.
- [13] J. Liao, "HeterSched: Structure-Aware Reinforcement Learning for Heterogeneous Distributed Task Scheduling," 2024.
- [14] Y. Yang, "Memory-Enhanced Transformer for Backend Microservice Anomaly Detection," 2024.
- [15] R. Meng, "Hierarchical Communication and Computation Scheduling for Efficient Federated Learning in Cloud-Edge Collaborative Intelligent Systems," 2024.
- [16] A. Aljohani, "Predictive analytics and machine learning for real-time supply chain risk mitigation and agility, " *Sustainability*, vol. 15, no. 20, p. 15088, 2023.
- [17] G. Baryannis, S. Validi, S. Dani, et al., "Supply chain risk management and artificial intelligence: state of the art and future research directions, " *International Journal of Production Research*, vol. 57, no. 7, pp. 2179-2202, 2019.
- [18] N. Rezki and M. Mansouri, "Machine learning for proactive supply chain risk management: Predicting delays and enhancing operational efficiency, " *Management Systems in Production Engineering*, vol. 32, no. 3, pp. 345-356, 2024.
- [19] G. Tuncel and G. Alpan, "Risk assessment and management for supply chain networks: A case study, " *Computers in Industry*, vol. 61, no. 3, pp. 250-259, 2010.

- [20] J. Um and N. Han, "Understanding the relationships between global supply chain risk and supply chain resilience: the role of mitigating strategies, " *Supply Chain Management: An International Journal*, vol. 26, no. 2, pp. 240-255, 2021.
- [21] M. N. Khan, P. Akhtar, L. L. Zhang, et al., "Operating in environments affected by uncertainty: Supply chain finance, timely information sharing using advanced technology, and financial performance in Supply Chain Management 4.0, " *Journal of General Management*, vol. 50, no. 1, pp. 37-52, 2024.
- [22] P. S. Kang and B. Bhawna, "Enhancing supply chain resilience through supervised machine learning: supplier performance analysis and risk profiling for a multi-class classification problem, " *Business Process Management Journal*, vol. 31, no. 7, pp. 2825-2848, 2025.