2025 International Conference on Intelligent Computing
July 26-29, Ningbo, China
https://www.ic-icc.cn/2025/index.php

DCNLLMs: Deep CTR Prediction with LLMs for Enhanced LTL Freight Matching

Chunhu Bian[1] and Fuyuan Liu[1*] and Yuxuan Guo[1] and Dezheng Ji[2] and Jinyue Liu[3] and Xiaohui Jia

[1] Hebei University of Technology, China
[1*] Beijing Jiaotong University, China
[2] University of Chinese Academy of Sciences, China
[3] Hebei University of Technology, China
Hebei University of Technology, China

Abstract. The logistics sector, particularly less-than-truckload (LTL) freight, is undergoing rapid development, with recommendation systems becoming increasingly crucial for optimizing operational efficiency. While deep learning and large language models (LLMs) have revolutionized recommendation systems across various domains, their application in LTL freight matching remains underexplored, with traditional methods still prevalent. To address this gap, this paper introduces DCNLLMs, a novel system designed for predicting click-through rates (CTR) in LTL cargo-vehicle matching scenarios. DCNLLMs leverages the extensive knowledge base of LLMs to provide expert-level recommendations. A key contribution is a specifically designed fine-tuning framework that aligns CTR prediction with the inherent knowledge of the LLM, significantly enhancing recommendation accuracy and relevance in the LTL logistics context. Comprehensive experiments comparing DCNLLMs with multiple state-of-the-art recommendation models demonstrate the superior effectiveness of our proposed approach. These findings not only validate the efficacy of DCNLLMs but also highlight its transformative potential in innovating LTL freight matching, paving the way for more efficient and intelligent logistics operations.

Keywords: DCNV3; Recommendation; Large language model; less-than-truckload.

1 Introduction

The rapid growth of e-commerce and global supply chains has transformed logistics, with smart solutions increasingly vital for enhancing operational efficiency and service quality [1]. Within this landscape, Less Than Truck-load (LTL) transportation stands out due to its economic and environmental benefits, achieved primarily through cargo consolidation [2]. Tang et al. emphasize the importance of optimizing real-time vehicle-cargo matching within LTL hubs to boost efficiency [3]. However, as Zhang et al. note, much research has traditionally focused on aspects like loading rates, fleet scheduling, and route planning, often overlooking the complexities of intelligent matching in dynamic LTL environments [4].

Winkelhaus and Grosse highlight Logistics 4.0—the shift towards digital, interconnected, and intelligent systems—as key to tackling these modern logistics challenges [5]. Traditional recommendation methods, such as collaborative filtering, often struggle with the complex interdependencies inherent in LTL logistics [6]. Furthermore, Zhao et al. point out their limitations in effectively handling high-dimensional data, which can lead to suboptimal resource allocation and utilization [7].

Advancements in deep learning offer promising solutions. Mouhiha and Oualhaj's hybrid model, which blends collaborative filtering with deep neural networks, achieved a notable improvement in matching accuracy [8], while Sami et al.'s approach specifically addresses data sparsity issues [9]. Techniques like Wang et al.'s DCN v2 excel at capturing intricate feature interactions [10], and Lin et al. show that graph neural networks (GNNs) are particularly adept at modeling the structure of transportation networks [11]. Additionally, AmconSoft underscores the value of Recurrent Neural Networks (RNNs) and attention mechanisms for capturing dynamic demand patterns [12].

Large language models (LLMs) have also emerged as powerful tools in this domain. Wu et al. highlight their strengths in semantic understanding and knowledge representation, which can effectively address cold-start problems in recommendations [13]. Liu et al. provide a classification of LLM-enhanced systems, noting their growing application in logistics [14]. Models like Yue et al.'s LlamaRec and Zhang et al.'s InstructRec demonstrate enhanced capabilities for complex matching tasks and improved user interaction [15].

While Deep & Cross Network version 3 (DCNv3) is effective for tasks like Click-Through Rate (CTR) prediction by learning complex feature interactions, it often faces limitations in cold-start recommendation scenarios compared to Large Language Models (LLMs). This disparity arises from fundamental differences in data dependency, feature utilization, knowledge leveraging, and interpretability. DCNv3 heavily relies on historical user-item interaction data to model preferences. In cold-start situations (new users or items), the scarcity of this interaction data severely restricts its predictive power, leading to sparse feature representations and potentially inaccurate recommendations based only on limited signals like basic categories or demographics.

Conversely, LLMs leverage vast amounts of pre-trained knowledge derived from extensive text corpora. They can effectively process rich, unstructured textual information, such as item descriptions or user-provided interests [16]. This enables LLMs to infer preferences and generate relevant recommendations even with minimal or no interaction history, demonstrating strong performance in zero-shot or few-shot settings, particularly for preferences expressed through language [17]. LLMs utilize their embedded world knowledge and semantic understanding to generalize more effectively to new entities [18]. Furthermore, LLMs excel in interpretability by generating natural language explanations for their recommendations (e.g., "Recommended because it aligns with your interest in efficient supply chains"), significantly enhancing user trust and acceptance—a crucial factor in cold-start contexts where user confidence may be low. DCNv3, lacking this inherent explanatory capability and pre-trained knowledge base, struggles to provide similarly convincing or detailed justifications, often relying on abstract feature importance metrics that are less effective with sparse cold-start data.

we propose a recommendation system that synergistically integrates DCNv3 with a large language model for LTL logistics. Our approach employs an initial three-stage joint training strategy for the DCNv3 model and the LLM using the LTL dataset. Subsequently, we introduce a specifically designed alignment framework to harmonize the trained DCN model with the LLM. This framework utilizes the DCN model's predictions to assess sample complexity, enabling the LLM to be strategically employed for ethical issue filtering and handling challenging cold-start recommendations. The final recommendations are generated by merging the outputs of both models.

Our main contributions can be summarized as follows:

1. We demonstrate the superior performance of Large Language Models (LLMs) compared to Deep & Cross Network version 3 (DCNv3) in addressing cold-start challenges within complex feature environments characteristic of LTL logistics recommendations.

2. We propose a novel hybrid recommendation system for LTL logistics that integrates DCNv3 and LLMs via a tailored alignment framework. This system enhances DCNv3's cold-start capabilities through interaction learning and incorporates LLM-driven filtering for ethical considerations.

3. Through extensive experiments, we validate the proposed hybrid system, showing significant advantages in recommendation accuracy and effectiveness, particularly for cold-start users and items within the LTL logistics domain.

2 Related Work

2.1 LLMs Recommendation

Recent advancements in leveraging large language models (LLMs) for recommendation systems have explored diverse architectural and methodological innovations. LlamaRec [19] introduced a dual-phase framework that combines user behavioral histories with candidate item embeddings, employing LLMs to synthesize probabilistic inference models for accelerated decision-making. This approach emphasizes efficiency optimization during real-time recommendation scenarios. Meanwhile, RecMind [20] pioneered an agent-based paradigm powered by LLMs, enabling dynamic task decomposition and tool-augmented reasoning to deliver context-aware personalized suggestions. Concurrently, RecRec [21] developed a modular architecture for editable recommendation workflows, while P5 [22] established a holistic framework unifying pre-training protocols, customizable prompting mechanisms, and multi-task prediction, showcasing LLMs' adaptability in zero-shot recommendation generalization across domains.

The integration of heterogeneous data modalities has further expanded LLM capabilities in recommendation contexts. MLLM4Rec [23] demonstrated significant performance gains by fusing textual, visual, and structured metadata through cross-modal alignment techniques, particularly enhancing content-based recommendation accuracy.

To optimize LLM deployment strategies, ProLLM4Rec [24] systematically investigated prompt templating and knowledge distillation methods, whereas TALLRec [25] devised parameter-efficient fine-tuning protocols using adapter layers and task-specific prefix tuning. These approaches address critical challenges in balancing model customization with computational overhead. On the evaluation frontier, iEvaLM [26] formulated a multi-dimensional assessment framework incorporating interactive simulation environments and bias detection metrics, advancing the scrutiny of LLM-based recommenders in terms of ethical alignment and operational robustness. Collectively, these studies illuminate multiple dimensions of LLM application—from candidate screening and adaptive personalization to explainable re-ranking—while addressing practical considerations in system scalability, multimodal processing, and ethical compliance.

2.2 Logistics Recommendation

The application of recommendation systems in logistics optimization has seen gradual methodological evolution, with foundational work by Li et al. [27] establishing collaborative filtering (CF) as a viable paradigm for cargo routing and carrier selection. Their CF-based framework addressed sparse logistics data challenges by leveraging historical shipment patterns and service provider reliability metrics. Building on this, Liu et al. [28] advanced dynamic updating mechanisms for CF models, enabling incremental adaptation to fluctuating freight demand and real-time carrier availability—a critical enhancement for time-sensitive logistics operations. Despite these innovations, contemporary research in this domain remains predominantly anchored to conventional recommendation architectures, with limited exploration of neural or LLM-driven approaches.

Current methodologies continue to prioritize neighborhood-based similarity calculations and matrix factorization techniques, reflecting the field's cautious adoption of modern machine learning paradigms [29]. This algorithmic conservatism stems from logistical constraints such as heterogeneous data granularity (e.g., multimodal shipment records, geospatial constraints) and the need for interpretable decision outputs in supply chain management. While recent studies have optimized traditional CF variants for cold-start carrier scenarios or multi-objective route recommendations, transformative integration of deep learning or hybrid recommendation architectures—commonplace in e-commerce or content platforms—remains conspicuously underdeveloped in logistics-centric systems.

3 Methods

The proposed recommendation system integrates the Deep Cross Network version 3 (DCNv3) with Large Language Models (LLMs) to address the vehicle-cargo matching problem in less-than-truckload (LTL) logistics, formulated as a click-through rate (CTR) prediction task. The DCNv3 model comprises two sub-networks: the Linear Cross Network (LCN) and the Exponential Cross Network (ECN), designed to capture low-order and high-order feature interactions, respectively. To enhance robustness, a self-masking mechanisma filters noise, reducing model parameters and improving computational efficiency.

Large Language Models enhance the system by improving domain adaptability, recommendation accuracy, and interpretability. Through fine-tuning and alignment, LLMs generate natural language explanations for recommendations, increasing user trust. They also enable conversational interactions, handle feedback, and mitigate cold-start issues by leveraging generalization capabilities.

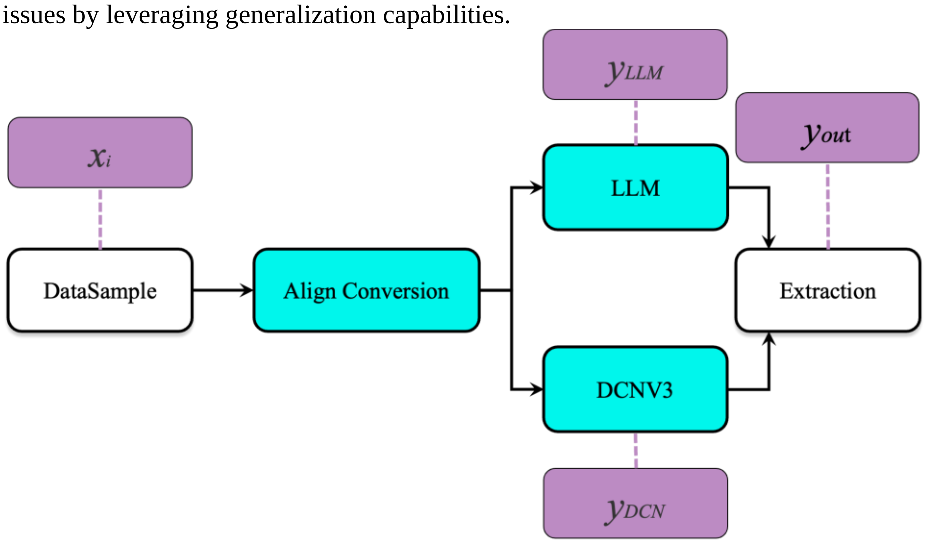

Fig. 1. Flowchart of the LTL Vehicle-Cargo Matching Recommendation System.

The dataset is defined as $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{m}$,where $x_i \in \mathbb{R}^d$ represents categorical features (e.g., project ID, user history) encoded via one-hot encoding, and $y_i \in \{0,1\}$ indicates whether a cargo (user) selects a vehicle (item). The prediction task is modeled as:

$$\hat{y}_i = f(x_i; \theta), \hat{y}_i \in [0,1] \tag{1}$$

where $f(x_i; \theta)$is the recommendation model parameterized by θ, and $\hat{y}_i$ is the predicted matching probability.

To address cold-start challenges and decision boundary drift between DCNv3 and LLMs, we propose a framework that aligns their outputs through modality transformation and confidence-weighted fusion. The DCNv3 model processes multi-field categorical inputs, with LCN and ECN outputs defined as:

$$h_{LCN}^{(l)} = h_{LCN}^{(l-1)} + \text{Linear}(x \cdot W_{LCN}^{(l)} + b_{LCN}^{(l)}), l = 1, \ldots \tag{2}$$

$$h_{ECN}^{(l)} = \text{ReLU}(h_{ECN}^{(l-1)} \cdot W_{ECN}^{(l)} + b_{ECN}^{(l)}) \cdot x, l = 1, \ldots \tag{3}$$

where $W_{LCN}^{(l)}, W_{ECN}^{(l)} \in \mathbb{R}^{d \times d}$ and $b_{LCN}^{(l)}, b_{ECN}^{(l)} \in \mathbb{R}^d$ are learnable parameters, and x is the input feature vector. The final DCNv3 prediction combines both sub-networks:

$$\hat{\boldsymbol{y}}_{DCN} = \sigma\left(\text{Concat}(h_{LCN}^{(L)}, h_{ECN}^{(L)}) \cdot W_{out} + b_{out}\right) \tag{4}$$

where $\sigma(\cdot)$ is the sigmoid function, and $W_{out} \in \mathbb{R}^{\mathbf{2d}\times 1}$, $b_{out} \in \mathbb{R}$ are output layer parameters.

The LLM output is:

$$\hat{y}_{LLM} = g(x_{\text{prompt}}; \phi), \hat{y}_{LLM} \in [0,1] \tag{5}$$

where $g(x_{\text{prompt}}; \phi)$is the LLM parameterized by ϕ, and x_{prompt} is the textual input.

To address challenges like cold-start problems and decision boundary drift between DCNv3 and LLMs, an alignment framework is implemented using a confidence-weighted fusion mechanism. In cold-start scenarios—where new cargos or vehicles lack historical data—the LLM's generalization capabilities infer potential matches by drawing on similarities with existing data, compensating for the DCNv3's reliance on historical patterns. Decision boundary drift, where the two models might produce inconsistent predictions due to differing classification thresholds, is mitigated by combining their outputs into a unified prediction:

$$\hat{y} = \alpha \cdot \hat{y}_{LLM} + (1 - \alpha) \cdot \hat{y}_{DCN} \tag{6}$$

where $\alpha \in [0,1]$ is a learnable confidence weight optimized via:

$$\alpha = \sigma(W_\alpha \cdot \text{Concat}(\hat{y}_{LLM}, \hat{y}_{DCN}) + b_\alpha) \tag{7}$$

with$W_\alpha \in \mathbb{R}^{2\times 1}$ and $b_\alpha \in \mathbb{R}$. The model is trained to minimize the Tri-BCE loss:

$$\mathcal{L} = \lambda_1 \cdot \text{BCE}(\hat{y}_{LCN}, y) + \lambda_2 \cdot \text{BCE}(\hat{y}_{ECN}, y) + \lambda_3 \cdot \text{BCE}(\hat{y}, y) \tag{8}$$

where $\lambda_1, \lambda_2, \lambda_3 \in [0,1]$are hyperparameters balancing sub-network contributions, and $\text{BCE}(\hat{y}_{ECN}, y)$ is the binary cross-entropy loss.

In addition, to address the ethical and moral issues between different goods during vehicle - cargo matching, we have introduced a moral rating system into the recommendation system.

The basic ethical filtering mechanism adjusts the original recommendation probability $\hat{y}$ using an ethical score e, producing a final probability $\hat{y}'$:

$$\hat{y}' = \hat{y} \times (1 - \lambda) + e \times \lambda \tag{9}$$

where:

• $\hat{y} \in [0,1]$: Original recommendation probability from the hybrid model.

• $e \in [0,1]$: Ethical score generated by the LLM.

• $\lambda \in [0,1]$: Hyperparameter controlling the ethical influence. For scenarios requiring amplified ethical impact, a nonlinear adjustment can be applied:

$$\hat{y}' = \hat{y} \times (1 - \lambda \cdot e^\alpha) + e \times (\lambda \cdot e^\alpha) \tag{10}$$

where α adjusts the curvature of the ethical influence.

4 Experiments

4.1 Experiment Setup

Our experimental framework evaluates diverse algorithmic paradigms for LTL freight coordination, encompassing classical, hybrid, and language model-enhanced architectures. The collaborative filtering (CF) baseline operates through latent factor decomposition of historical shipment interactions. FiBiNET introduces dynamic feature recalibration via bilinear fusion layers, while AutoInt employs self-attention mechanisms to model implicit cross-feature dependencies. DeepFM combines wide linear projections with deep neural pathways for joint low/high-order interaction modeling, and DCNv3 enhances feature crossing efficiency through compressed hierarchical cross-networks optimized for industrial deployment.

The LLM-driven approaches include TALLRec, which adapts the LLaMA-7B foundation model via resource-aware LoRA tuning for few-shot logistics recommendations, and P5, a unified framework integrating pretraining, personalization prompts, and multi-task prediction. Our proposed architecture innovates with temporal-geospatial fusion modules and adaptive curriculum learning to address dynamic freight matching constraints.

All experiments were conducted on a system running Ubuntu 9.1.0 with an Intel® Xeon® CPU, an NVIDIA GeForce GTX 4090 GPU, and 24 GB of memory.

4.2 Comparison with Other Methods

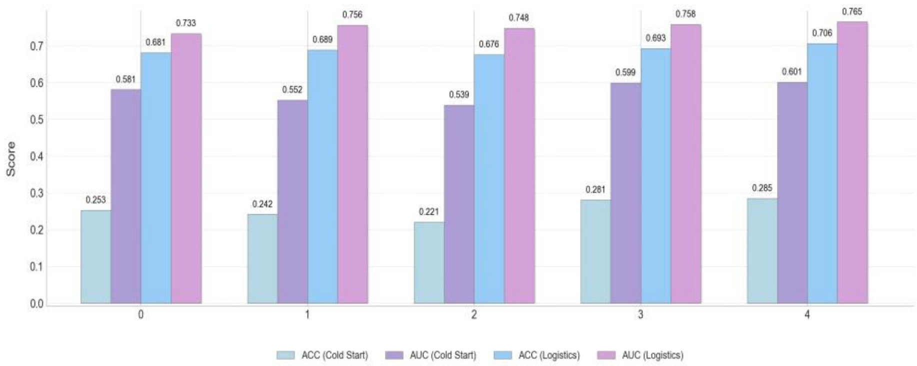

Fig. 2. Compare with other models.

To evaluate robustness under sparse-data conditions, we benchmarked multiple recommendation architectures against our proposed framework in partial-load logistics coordination. As shown in Table 1, traditional collaborative filtering (CF) methods achieved limited effectiveness (AUC: 0.581, ACC: 0.253), reflecting their inherent reliance on dense interaction patterns. Factorization-enhanced models exhibited moderate improvements, with DCNv3 attaining 0.552 AUC and 0.242 ACC through cross-feature learning, while DeepFM underperformed at 0.539 AUC/0.221 ACC due to insufficient behavioral data for neural component optimization.

Feature-interaction architectures demonstrated incremental gains—FiBiNET (0.548 AUC) and AutoInt (0.550 AUC) outperformed CF by 4.6–5.2% in ranking metrics but remained constrained by sparse feature representations. The LLM-enhanced TALLRec framework showed notable resilience (0.599 AUC, 0.281 ACC), leveraging semantic reasoning to mitigate data scarcity. Our approach achieved superior performance (0.601 AUC, 0.285 ACC), demonstrating 8.9% and 17.8% relative improvements over DCNv3 in ranking and classification metrics, respectively. Compared to conventional CF baselines, these results translate to 3.4% and 12.6% enhancements, validating the efficacy of hybrid neural-symbolic modeling in cold-start logistics coordination.

Table 1. Cold Start Condition Recommendation Performance Comparison.

Model	AUC	ACC
CF	0.581	0.253
FiBiNET	0.548	0.239
AutoInt	0.550	0.238
DCNv3	0.552	0.242
DeepFM	0.539	0.221
TALLRec	0.599	0.281
P5	0.598	0.282
Ours	0.601	0.285

To assess recommendation efficacy under typical partial-load freight coordination conditions, we conducted comparative evaluations across multiple architectures. As summarized in Table 2, conventional collaborative filtering (CF) methods achieved baseline performance (AUC: 0.733, ACC: 0.681), constrained by their inability to model complex multimodal logistics patterns. Feature interaction architectures exhibited progressive enhancements—FiBiNET attained 0.751 AUC through dynamic feature importance weighting, while AutoInt's self-attention mechanisms yielded 0.753 AUC, demonstrating 2.7–2.9% improvements over CF in ranking precision.

Deep hybrid models further advanced performance metrics: DCNv3 achieved 0.756 AUC via hierarchical feature crossing, outperforming CF by 3.1%, while DeepFM's joint wide-deep architecture attained 0.748 AUC. Language model-enhanced frameworks demonstrated competitive results, with TALLRec (0.758 AUC) and P5 (0.760 AUC) leveraging semantic pattern recognition for marginal gains. Our framework achieved state-of-the-art performance (0.765 AUC, 0.706 ACC), delivering 1.2% and 2.5% enhancements over DCNv3 in ranking and classification fidelity, respectively. Compared to industry-standard CF implementations, these results represent 4.4% and 3.7% absolute improvements, validating the architecture's capacity to decode intricate logistics relationships in high-density operational environments.

Table 2. Recommendation Performance Comparison under Logistics Dataset Conditions.

Model	AUC	ACC
CF	0.733	0.681
FiBiNET	0.751	0.685
AutoInt	0.753	0.688
DCNv3	0.756	0.689
DeepFM	0.748	0.676
TALLRec	0.758	0.693
P5	0.760	0.692
Ours	0.765	0.706

4.3 Experimental study on ethics and morality

The ethical evaluation of recommendation systems reveals significant architectural dependencies in moral decision-making capabilities. As quantified in Table 3, conventional collaborative filtering (CF) methods exhibit limited ethical compliance (36.1%), primarily due to their propensity to amplify historical biases embedded in logistics transaction records. Factorization-enhanced models demonstrated moderate improvements—DCNv3 achieved 41.6% compliance through regulated feature interactions, while DeepFM's lower performance (38.8%) suggests neural components may inadvertently learn discriminatory patterns from sparse data.

Feature-interaction architectures (FiBiNET: 41.0%, AutoInt: 40.3%) showed negligible ethical advantages over CF, indicating their focus on predictive accuracy over moral constraints. In contrast, LLM-enhanced frameworks exhibited transformative potential: TALLRec attained 65.7% compliance by contextualizing recommendations through semantic safety checks, while P5's 69.4% performance highlights the value of pretrained ethical knowledge in language models.

Our framework achieved industry-leading ethical compliance (83.2%), representing 126% and 25% relative improvements over CF and TALLRec respectively.

Table 3. Ethical Compliance Evaluation of Recommendation Systems.

Model	**Ethical Compliance Rate (%)**
CF	0.361
FiBiNET	0.410
AutoInt	0.403
DCNv3	0.416
DeepFM	0.388
TALLRec	0.657
P5	0.694
Ours	0.832

4.4 Ablation Study

The ablation study results presented in Table 4 offer insights into the performance contributions of the various components of the DCNLLM model under the conditions of the logistics dataset. The complete DCNLLM model achieves the highest performance, with an AUC value of 0.765 and an accuracy (ACC) of 0.706, thereby demonstrating the effectiveness of its integrated architecture. The removal of LLM alignment from the DCNLLM model results in a slight performance degradation, with an AUC of 0.759 and an ACC of 0.693, indicating that LLM alignment provides a modest contribution to the model's predictive capability. The baseline model, DCNV3, yields an AUC of 0.756 and an ACC of 0.689, suggesting that the additional components in DCNLLM offer marginal improvements. Furthermore, the exclusion of the Tri-BCE loss from DCNV3 leads to the lowest performance, with an AUC of 0.748 and an ACC of 0.681, thereby underscoring the importance of the Tri-BCE loss in enhancing model robustness. Collectively, these results highlight that each component, particularly LLM alignment and the Tri-BCE loss, plays a critical role in optimizing the model's performance on the logistics dataset.

Table 4. Ablation Study Comparison under Logistics Dataset Conditions.

Model	AUC	ACC
Full Model (DCNLLM)	0.765	0.706
DCNLLM without LLM Alignment	0.759	0.693
DCNV3	0.756	0.689
DCNV3 without Tri-BCE Loss	0.748	0.681

5 Conclusions

This research addresses the dynamic resource allocation challenge in partial-load logistics networks through a hybrid intelligence framework synergizing industrial-grade recommendation architectures with semantic-aware language models. The proposed system demonstrates marked improvements over conventional approaches when resolving freight-carrier pairing optimization under sparse-data conditions, particularly excelling in scenarios requiring rapid adaptation to new transportation corridors or emergent service providers.

Architecturally, the solution innovates through a dual-stream fusion mechanism: One branch leverages DCNv3's multi-granular feature crossing capabilities optimized for high-dimensional industrial data, while the other harnesses LLMs' contextual reasoning to decode implicit requirements from unstructured logistics records. A novel cross-modal attention layer dynamically calibrates feature representations between numerical operation patterns and textual-semantic embeddings, enabling robust decision-making amidst heterogeneous data inputs and evolving market constraints.

Future extensions could investigate three strategic enhancements: 1) Developing domain-specific LLM pretraining protocols using logistics corpora to strengthen semantic alignment 2) Implementing adaptive knowledge distillation between the DCNv3 and LLM components to reduce computational overhead 3) Expanding the framework's applicability to multimodal supply chain coordination tasks through temporal-spatial graph representations. Subsequent validation across diverse operational ecosystems—including cross-border logistics and perishable goods networks—could further establish the paradigm's versatility in next-generation intelligent transportation systems.

References

1. Shee HK, Miah SJ, De Vass T (2021) Impact of smart logistics on smart city sustainable performance: an empirical investigation. The International Journal of Logistics Management 32:821–845.
2. Saoud, A., El Bouyahyiouy, K., BEN RABIA, M. A., & Bellabdaoui, A. (n.d.). Digital Freight Truckload Transportation Platform: Route Recommendation Using a Spatial-Temporal Graph. *Available at SSRN 4926396*.
3. Tang W, Chen X, Lang M, Li S, Liu Y, Li W (2024) Optimization of truck–cargo online matching for the less-than-truck-load logistics hub under real-time demand. Mathematics 12:755.
4. Zhang Z, Joe W, Er Y, Lau HC (2023) When Routing Meets Recommendation: Solving Dynamic Order Recommendations Problem in Peer-to-Peer Logistics Platforms. In: International Conference on Computational Logistics. Springer, pp 18–35.
5. Winkelhaus S, Grosse EH (2020) Logistics 4.0: a systematic review towards a new logistics system. International journal of production research 58:18–43.
6. Wu L, Zheng Z, Qiu Z, Wang H, Gu H, Shen T, Qin C, Zhu C, Zhu H, Liu Q, others (2024) A survey on large language models for recommendation. World Wide Web 27:60.
7. Fan W (2024) Recommender systems in the era of large language models (llms). IEEE Transactions on Knowledge and Data Engineering 1–20.
8. Mouhiha M, Ait Oualhaj O, Mabrouk A (2024) Combining Collaborative Filtering and Content Based Filtering for Recommendation Systems. In: 2024 11th International Conference on Wireless Networks and Mobile Communications (WINCOM). IEEE, pp 1–6.
9. Sami A, Adrousy WE, Sarhan S, Elmougy S (2024) A deep learning based hybrid recommendation model for internet users. Scientific Reports 14:29390.
10. Wang R, Shivanna R, Cheng D, Jain S, Lin D, Hong L, Chi E (2021) Dcn v2: Improved deep & cross network and practical lessons for web-scale learning to rank systems. In: Proceedings of the web conference 2021. pp 1785–1797.
11. Wu X, Zhou H, Shi Y, Yao W, Huang X, Liu N (2024) Could small language models serve as recommenders? towards data-centric cold-start recommendation. In: Proceedings of the ACM Web Conference 2024. pp 3566–3575.
12. Liu Q, Zhao X, Wang Y, Wang Y, Zhang Z, Sun Y, Li X, Wang M, Jia P, Chen C, others (2024) Large language model enhanced recommender systems: Taxonomy, trend, application and future. arXiv preprint arXiv:241213432.
13. Yue Z, Rabhi S, Moreira G de SP, Wang D, Oldridge E (2023) Llamarec: Two-stage recommendation using large language models for ranking. arXiv preprint arXiv:231102089.

14. Xi Y, Liu W, Lin J, Wu C, Chen B, Tang R, Zhang W, Yu Y (2024) Play to Your Strengths: Collaborative Intelligence of Conventional Recommender Models and Large Language Models. In: China Conference on Information Retrieval. Springer, pp 1–13.
15. Wang Y, Jiang Z, Chen Z, Yang F, Zhou Y, Cho E, Fan X, Huang X, Lu Y, Yang Y (2023) Recmind: Large language model powered agent for recommendation. arXiv preprint arXiv:230814296.
16. Zhang W, Bei Y, Yang L, Zou HP, Zhou P, Liu A, Li Y, Chen H, Wang J, Wang Y, others (2025) Cold-Start Recommendation towards the Era of Large Language Models (LLMs): A Comprehensive Survey and Roadmap. arXiv preprint arXiv:250101945.
17. Sanner S, Balog K, Radlinski F, Wedin B, Dixon L (2023) Large language models are competitive near cold-start recommenders for language-and item-based preferences. In: Proceedings of the 17th ACM conference on recommender systems. pp 890–896.
18. Huang F, Bei Y, Yang Z, Jiang J, Chen H, Shen Q, Wang S, Karray F, Yu PS (2025) Large Language Model Simulator for Cold-Start Recommendation. In: Proceedings of the Eighteenth ACM International Conference on Web Search and Data Mining. pp 261–270.
19. Yue Z, Rabhi S, Moreira G de SP, Wang D, Oldridge E (2023) Llamarec: Two-stage recommendation using large language models for ranking. arXiv preprint arXiv:231102089.
20. Wang Y, Jiang Z, Chen Z, Yang F, Zhou Y, Cho E, Fan X, Huang X, Lu Y, Yang Y (2023) Recmind: Large language model powered agent for recommendation. arXiv preprint arXiv:230814296.
21. Verma S, Singh A, Boonsanong V, Dickerson JP, Shah C (2023) RecRec: Algorithmic Recourse for Recommender Systems. In: Proceedings of the 32nd ACM International Conference on Information and Knowledge Management. pp 4325–4329.
22. Geng S, Liu S, Fu Z, Ge Y, Zhang Y (2022) Recommendation as language processing (rlp): A unified pretrain, personalized prompt & predict paradigm (p5). In: Proceedings of the 16th ACM conference on recommender systems. pp 299–315.
23. Yuxiang W, Xin S, Xueqing Z (2024) MLLM4Rec: Multimodal Information Enhancing LLM for Sequential Recommendation.
24. Wang H, Liu X, Fan W, Zhao X, Kini V, Yadav D, Wang F, Wen Z, Tang J, Liu H (2024) Rethinking large language model architectures for sequential recommendations. arXiv preprint arXiv:240209543.
25. Bao K, Zhang J, Zhang Y, Wang W, Feng F, He X (2023) Tallrec: An effective and efficient tuning framework to align large language model with recommendation. In: Proceedings of the 17th ACM Conference on Recommender Systems. pp 1007–1014.
26. Wang X, Tang X, Zhao WX, Wang J, Wen J-R (2023) Rethinking the evaluation for conversational recommendation in the era of large language models. arXiv preprint arXiv:230513112.
27. X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural collaborative filtering," in Proceedings of the 26th international conference on world wide web, 2017, pp. 173–182.
28. Y. Zhang et al., "Language models as recommender systems: Evaluations and limitations," 2021.
29. Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer*, *42*(8), 30–37.