Journal of Engineering, Project, and Production Management
2026, 16(7), 2025-309

DLAGNN: A Dual-Layer Attention Graph Neural Network Framework with Dynamic Feature Fusion and Multi-Task Learning for Customer Relationship Management in E-Commerce

Junyi Yang
Full-time Faculty, School of Artificial Intelligence, Chongqing Jianzhu College, Chongqing,400072, China, E-mail: y.yjunyi@outlook.com

Project Management
Received December 8, 2025; revised January 22, 2026; accepted July 2, 2026
Available online July 20, 2026

Abstract: The customer-centric transformation of e-commerce makes it difficult for traditional customer relationship management to meet the needs of enterprises to accurately tap into customer value. Therefore, the research aims to build an intelligent model that integrates multi-source data, accurately predicts customer needs, and guides personalized marketing. The study utilizes the Alibaba Tianchi dataset, comprising 100,000 user interaction logs, to validate the proposed Dual-Layer Attention Graph Neural Network (DLAGNN) framework. Distinct labels were constructed for multi-task learning: customer churn was defined by a 30-day inactivity threshold, while Customer Lifetime Value (CLV) was calculated from historical transaction amounts. The framework employs a dual-layer attention mechanism to capture long-term and short-term preferences, respectively, integrates dynamic features such as timeliness and profitability, and simultaneously optimizes preference prediction, churn warning, and CLV prediction. Results indicate that the model converged efficiently and achieved superior predictive accuracy compared to baseline models such as GNN-SR and TA-RNN. Its customer churn warning could be issued significantly in advance of the actual event, extending the retention window. In the simulation test, the Return on Investment (ROI) in churn intervention and CLV improvement demonstrated the model's practical effectiveness. This model effectively solves data sparsity and model commercial adaptability, providing a feasible path for applying deep learning in Customer Relationship Management (CRM).

Keywords: E-commerce personalized marketing; customer relationship management; data mining; graph neural network; dynamic feature fusion.

DOI 10.32738/JEPPM-2025-308

---

1. Introduction

Driven by big data, cloud computing and artificial intelligence technology, the e-commerce industry is undergoing a profound change, that is, from the traditional product-centered marketing model to customer-centered personalized services (Groumpos, 2023). In this context, the connotation and extension of Customer Relationship Management (CRM) have also expanded. Traditional passive customer service can no longer meet enterprises needs to deeply explore customer value amid fierce competition. Massive amounts of customer behavior, transaction and interaction data provide enterprises with unprecedented insights, enabling them to proactively and accurately predict customer needs and formulate personalized marketing strategies. However, in practice, it is not easy to transform these multi-source and heterogeneous dynamic data into actionable business insights. Data mining, as a technology for discovering meaningful patterns and regularities in massive amounts of data, provides strong support for solving this challenge. Traditional CRM methods, such as customer-segmentation-based marketing strategies, often struggle to capture the subtle preferences and instantaneous changes in demand of individual customers, leading to a lack of targeted marketing activities and an inability to effectively improve customer satisfaction and loyalty (Alves et al., 2023). However, existing data analysis technologies still face challenges such as data sparsity, dynamic preference drift, and business goal fusion when dealing with complex, dynamic user-product interaction relationships (Zhang et al., 2023). These challenges limit enterprises' ability to translate data mining insights into business value, making it difficult to allocate marketing resources accurately, provide early warnings of customer churn risk, and maximize Customer Lifetime Value (CLV) (Mao et al., 2023). Therefore, building an intelligent framework that integrates multi-source data, accurately predicts customer needs, and effectively guides personalized marketing has become a key challenge.

Existing research focuses on e-commerce personalized marketing demand prediction and CRM, exploring topics ranging from multi-source data integration and prediction model design to customer management strategy optimization and other dimensions, yielding various theoretical results and practical solutions. Raji et al. (2024) pointed out that artificial intelligence-driven personalization technology in e-commerce could significantly improve customer engagement and loyalty, give rise to trends such as chatbot interaction, predictive analysis and inventory optimization and promote the transformation of the online retail ecosystem. Mykhalchenko and Tytarenko, (2023). used a mixed method to rank e-commerce data analysis tools, confirming that correct analysis tools could effectively avoid financial losses. Ferrer-Estévez et al. (2023)systematically combed the existing knowledge in the field of CRM to determine its main research schools and future agenda. 139 documents were analyzed using the PRISMA methodology, content analysis and bibliometric tools. The research results developed a classification framework with seven categories, identified the main contributors and findings in this field, and provided clear guidance for future research, highlighting the importance of CRM in sustainable development. The enhanced Conformer model designed by Lin et al. could reduce errors in online ride-hailing demand prediction, and its feature fusion approach provided a reference for personalized demand prediction in e-commerce. Zhao et al. (2022) used radio-frequency identification data to improve customer segmentation accuracy and identify customer behavior patterns. They integrated radio-frequency identification data with traditional point-of-sale data and built a new classification model based on a sequence-to-sequence learning architecture. This model showed higher accuracy and area under the curve in customer classification and identification tasks. Experimental results showed that its performance was superior to that of other benchmark models and that the behavioral description variables were effective across heterogeneous customer groups.

Graph Neural Networks (GNNs), as deep learning models that are effective at processing non-Euclidean data, have been widely used in social network analysis, intelligent recommendations, traffic demand prediction, and other fields. It is especially outstanding at capturing complex relationships among entities. It modeled entities as nodes and associations as edges, and combined node information aggregation to mine potential associations. It had significant advantages in solving data sparsity and improving dynamic prediction accuracy, and provided effective technical support for exploring complex problems in multiple fields (Gao et al., 2023). Aiming at the demand prediction for e-commerce personalized marketing, Rahmani et al. (2023) systematically reviewed the extensive applications of GNNs in intelligent transportation, particularly their evolution and latest progress in demand prediction. The results showed that the GNN framework not only demonstrated superior performance but also revealed numerous emerging research opportunities, including edge computing, multimodal models and the integration of unsupervised and reinforcement learning methods, providing valuable inspiration for the future development of personalized marketing demand prediction models. In response to the uncertainty quantification in e-commerce personalized marketing demand prediction, Wang et al. (2024a) applied the probabilistic GNN framework to the transportation demand prediction task from a cross-domain application perspective. This framework was to integrate deterministic and probabilistic assumptions to predict the spatial and temporal uncertainty of demand. The results showed that probabilistic assumptions were crucial to prediction accuracy, with truncated Gaussian and Laplace distributions performing best. Even with significant domain shifts, probabilistic GNNs can still stabilize prediction uncertainty and reveal spatiotemporal distribution patterns, providing an important reference for building more resilient prediction models in the e-commerce field. In response to the question of how to use social relationships to improve prediction accuracy in e-commerce personalized marketing demand prediction, Sharma et al. (2024) systematically reviewed the application of GNNs in this field from the perspective of social recommendation systems. They followed the PRISMA framework, carefully classified 84 high-quality documents, and proposed an innovative classification system based on input type and GNN architecture. The results showed that GNN could effectively capture social influence among users, significantly enhance understanding of user preferences, and provide a new perspective and rich technical options for personalized marketing demand prediction in the e-commerce field.

To sum up, existing research has made some progress in artificial intelligence-driven personalized marketing, e-commerce data analysis tool selection, and GNN cross-domain prediction applications. However, it has not yet effectively integrated customers long- and short-term preferences to address dynamic demand drift, and it lacks a deep coupling between business goals and prediction models. To address the limitations of existing approaches and bridge the gap between technical prediction and business application, this study focuses on the following three research questions (RQs): RQ1: How can dynamic variations in user preferences be effectively captured from sparse and heterogeneous e-commerce interaction data? RQ2: To what extent does the integration of dual-layer attention mechanisms improve the representation learning of user behaviors compared to static graph models? RQ3: How can the trade-off between churn prediction accuracy and CLV estimation precision be balanced within a unified Multi-Task Learning (MTL) framework? In view of the above shortcomings, the research uses data mining technology as the core framework and first relies on narrow data mining methods, such as association rules and clustering, to mine basic rules from multi-source data. A dual-layer attention-mechanism GNN model (a deep learning model in the broad scope of data mining) that combines Dynamic Feature Fusion (DFF) with an MTL strategy is proposed. By capturing long- and short-term preferences in layers and incorporating timeliness and profitability characteristics, it specifically addresses data sparsity and model business adaptability. This study aims to overcome the limitations of traditional CRM passive analysis, achieve seamless integration of accurate customer demand prediction with personalized marketing, and provide scientific tools for e-commerce companies to optimize marketing resource allocation, reduce customer churn, and advance customer-centric marketing models. The innovation of the research lies in constructing a “long-term preference-short-term dynamics” dual-layer attention GNN framework, using node self-attention to address long-term interest modeling, and using time-aware attention to capture demand drift. The MTL is introduced to simultaneously optimize preference, churn and CLV prediction, integrating profit characteristics to balance customer experience and corporate income and providing a new technical solution for e-commerce personalized CRM that is both accurate and practical.

2. Methods and Materials

2.1. Customer Demand Prediction and Behavioral Pattern Recognition Model Construction

While traditional data mining methods, such as matrix decomposition and collaborative filtering, provide a basis for interaction analysis, they often struggle to capture dynamic shifts in preferences and are strictly limited by data sparsity (Wang et al., 2023). To address these limitations, this study utilizes GNN to construct the core model architecture. By transforming customers, products, and attributes into graph structures, the proposed model captures complex interaction logic beyond the capability of conventional linear models. The relevant model architecture is shown in Fig. 1, which can provide a global view for subsequent mechanism analysis.

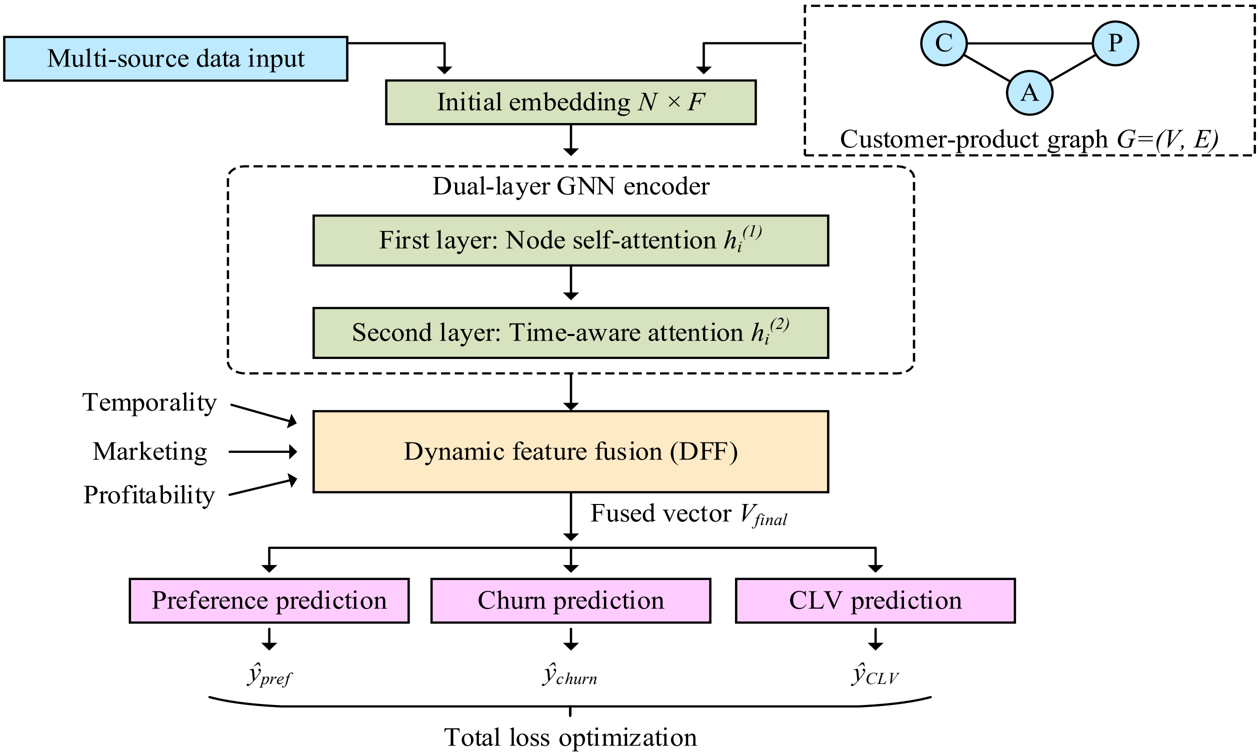


Fig. 1. Dual-layer attention mechanism GNN architecture diagram

As illustrated in Fig. 1, the DLAGNN framework processes data in a hierarchical manner. First, multi-source heterogeneous data are structured into a customer-product graph, denoted as $G = (V, E)$, converting massive customer behavior, product, and marketing data into a graph structure (Zhong et al., 2023). This structure is then transformed into an initial embedding matrix of size $N \times F$. Subsequently, the dual-layer GNN encoder extracts long-term and short-term preference representations, denoted as $h_i^{(1)}$ and $h_i^{(2)}$, respectively. The DFF module integrates these embeddings with external features (temporality, marketing, profitability) to generate a final fused vector V_{final}. Finally, the model utilizes MTL to output predictions for preference ($\hat{y}_{pref}$), churn ($\hat{y}_{churn}$), and CLV ($\hat{y}_{CLV}$) simultaneously, optimizing the total loss function. This framework lays the foundation for improved prediction accuracy. To illustrate the data transformation process intuitively, a specific example is shown in Fig. 2.

As shown in Fig. 2, the customer-product-attribute relationship map centers on customers and connects all relevant information through their behavioral trajectories. The graph first models the interaction between customers and products as behavioral relationship edges connecting them, with dynamic information such as timestamps attached, which intuitively reflects customers' behavior patterns (Li et al., 2024b; Yao et al., 2023b). Subsequently, to enhance information richness, the graph introduced additional attribute nodes, including product attributes and marketing events. These nodes are connected to customers or products through relationship edges, achieving a unified representation of the static attributes of goods and the external dynamic environment (Berahmand et al., 2024). Ultimately, this multi-dimensional and multi-type graph structure provides high-quality input data for the dual-layer attention mechanism GNN model to comprehensively understand customer needs. The initial embedding of the customer node is shown in Eq. (1).

$$\boldsymbol{u}_c = \frac{1}{|S_c|}\sum_{p \in S_c} \boldsymbol{e}_p \tag{1}$$

In Eq. (1), $\boldsymbol{u}_c$ represents the initial node embedding vector of customer c. S_c denotes the set of products that customer c has historically interacted with. $|S_c|$ is the total number of products in this set. $\boldsymbol{e}_p$ represents the embedding vector of product p, which aggregates the product's static attributes and dynamic characteristics. The customer-product interaction edge weight is shown in Eq. (2).

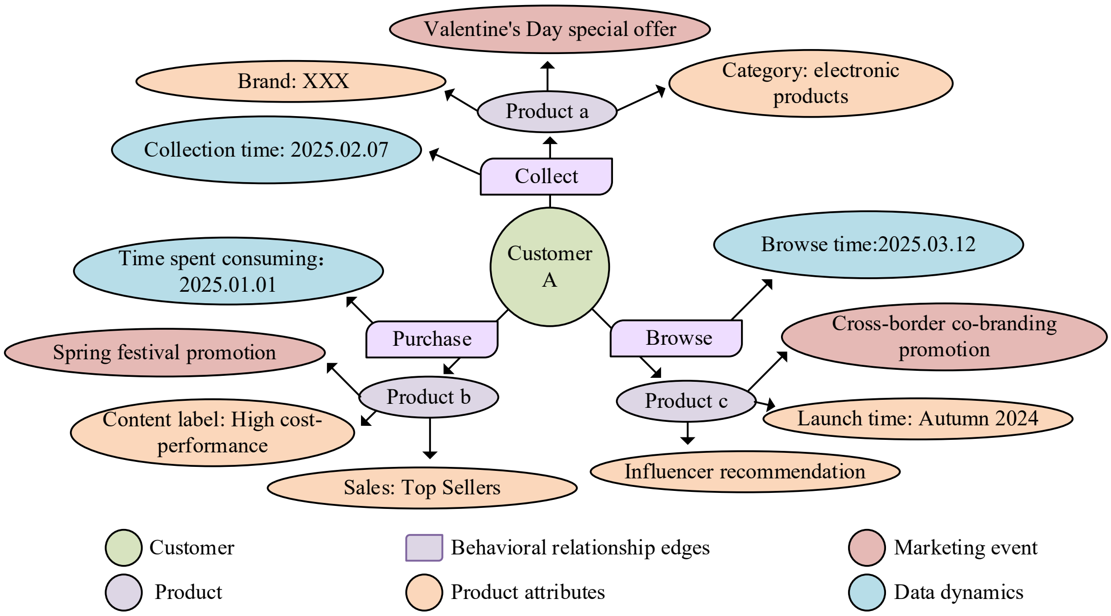


Fig. 2. Example of customer-product-attribute relationship map

$$w_{c,p} = \omega \cdot \eta^{\Delta t} \tag{2}$$

In Eq. (2), $w_{c,p}$ is the calculated weight of the interaction edge between customer c and product p. ω is the weight coefficient corresponding to the interaction type (e.g., browsing, purchasing). η is the time attenuation coefficient (consistent with the DFF module in Eq. 7). Δt signifies the time difference (in days) between the interaction event and the current timestamp. The product-attribute association weight is shown in Eq. (3).

$$v_{p,a} = \frac{N_{p,a}}{M_p} \tag{3}$$

In Eq. (3), $v_{p,a}$ represents the weight of the edge connecting product p and attribute a. $N_{p,a}$ signifies the frequency with which product p is associated with attribute a (e.g., appearing in a specific marketing campaign). M_p is the total number of attributes associated with product p. To specifically explain how the model uses this map to perform deep feature learning, the research will further focus on the first-layer attention mechanism. The detailed process of customer long-term preference modeling is shown in Fig. 3.

As shown in Fig. 3, the customer long-term preference modeling process is centered on node self-attention aggregation, which aims to extract stable and persistent interest features from customers' complex historical interactions. The model first traverses the customer's historical interaction records to collect all the products they have interacted with. Then, for each product node that the customer node has interacted with, the model will calculate the attention weight. This step can automatically identify products that contribute highly to the customer's long-term preference (Yao et al., 2023a). The customer-product attention weight is shown in Eq. (4).

$$\alpha_{c,p} = \frac{exp\left(\frac{(u_c)^\top e_p}{\tau}\right)}{\sum_{k \in S_c} exp\left(\frac{(u_c)^\top e_k}{\tau}\right)} \tag{4}$$

In Eq. (4), $\alpha_{c,p}$ denotes the normalized attention weight of customer c towards product p. $(\boldsymbol{u}_c)^\top \boldsymbol{e}_p$ represents the dot product between the customer vector and product vector. τ is the temperature coefficient used to smooth the attention distribution. The $softmax$ function ensures that the sum of attention weights equals 1. Based on these weights, the model will aggregate neighbor information, perform a weighted sum of the embedding vectors of related items, and generate a vector representing the customer's aggregated behavior. The weighted aggregation of neighbor information is shown in Eq. (5).

$$\boldsymbol{v}_c = \sum_{p \in S_c} \alpha_{c,p} \cdot \boldsymbol{e}_p \tag{5}$$

In Eq. (5), $\boldsymbol{v}_c$ is the aggregated neighbor information vector for customer c. It represents the weighted sum of the embedding vectors $\boldsymbol{e}_p$ of all interacted products, weighted by their respective attention scores $\alpha_{c,p}$. Finally, based on nonlinear transformation, an embedding vector representation that can reflect the long-term interests of customers is obtained. The long-term preference nonlinear transformation is shown in Eq. (6).

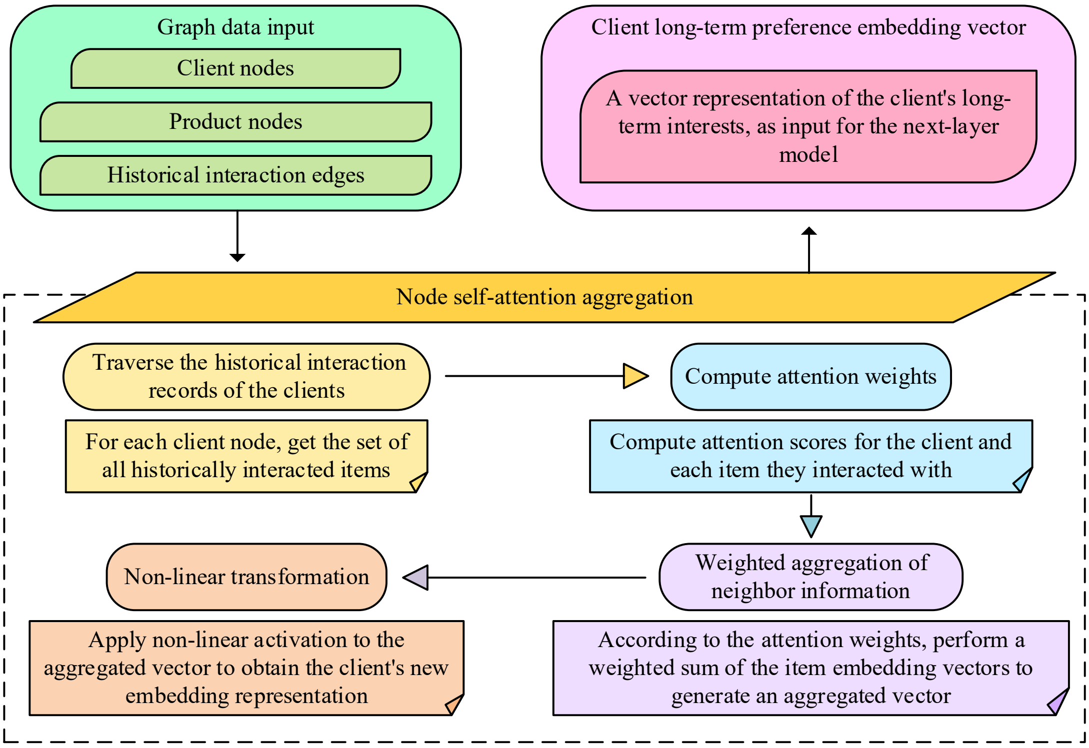


Fig. 3. Customer long-term preference modeling flow chart

$$\boldsymbol{h}_c^{(1)} = \mathrm{ReLU}(\mathbf{W}_1 \cdot \boldsymbol{v}_c + \boldsymbol{b}_1) \tag{6}$$

In Eq. (6), $\boldsymbol{h}_c^{(1)}$ represents the long-term preference embedding vector of customer c (corresponding to the "First layer" output in Fig. 1). $\boldsymbol{W}_1$ and $\boldsymbol{b}_1$ signify the learnable weight matrix and bias vector, respectively. $\mathrm{ReLU}(\cdot)$ signifies the non-linear activation function. This vector serves as the input for the subsequent second-layer attention mechanism. This vector is used as the output of the first-layer GNN, which effectively solves the data sparsity problem and generates a representation rich in semantic information for each customer. This long-term preference vector will serve as input to the next layer. The second-layer attention mechanism focuses on capturing customers' short-term dynamic preferences, modeling recent behavioral sequences through time-aware attention, and incorporating contextual information, such as external marketing activities, to capture and predict customer preference drift in real time, thereby comprehensively identifying customer behavior patterns (Wang et al., 2024b).

2.2. DFF and Model Optimization Strategies

The dual-layer attention-mechanism GNN model, constructed within the data-mining framework, effectively captures customers' long- and short-term preferences. However, it still faces challenges in integrating dynamic scene factors and collaboratively optimizing multiple business goals. As key strategies to improve model adaptability and practical value, DFF and MTL are introduced to account for data timeliness and balance conflicting business objectives (Chen et al., 2024; Xu et al., 2023). The proposed DFF mechanism employs a parallel processing architecture to independently encode dynamic characteristics, specifically timeliness, external marketing, and profitability, before fusing them with the node embeddings generated by the GNN. The detailed optimization workflow is illustrated in Fig. 4.

As shown in Fig. 4, the DFF mechanism operates in a "Left-to-Right" flow, transforming raw multi-source data into a unified representation through specific encoders. The process begins with four parallel input streams: the GNN node embedding (denoted as $U_L||p$, raw timeliness data T_{raw}, raw marketing data M_{raw}, and raw profitability data P_{raw}. First, the GNN node embedding, which contains the learned user preference information, undergoes an identity mapping to preserve its structural integrity. Simultaneously, the dynamic features are processed through dedicated encoders to normalize their varying scales and distributions (Hou et al., 2023). For timeliness features, a time-decay function is applied to capture the diminishing relevance of historical interactions. The encoding of the timeliness feature is shown in Eq. (7).

$$E_T = \eta^{1/\Delta t} \times T_{raw} \tag{7}$$

In Eq. (7), E_T denotes the encoded timeliness feature vector. T_{raw} represents the raw timeliness input. η is the timeliness attenuation coefficient (set to $0 < \eta < 1$, with a research value of $\eta = 0.85$). Δt signifies the time difference (in days) between the current timestamp and the latest interaction event. This encoding ensures that recent behaviors exert a stronger influence on the model's predictions. For external marketing features, such as promotion intensity, Min-Max normalization is used to map the data to the standard [0,1] interval. The encoded external marketing feature is shown in Eq. (8).

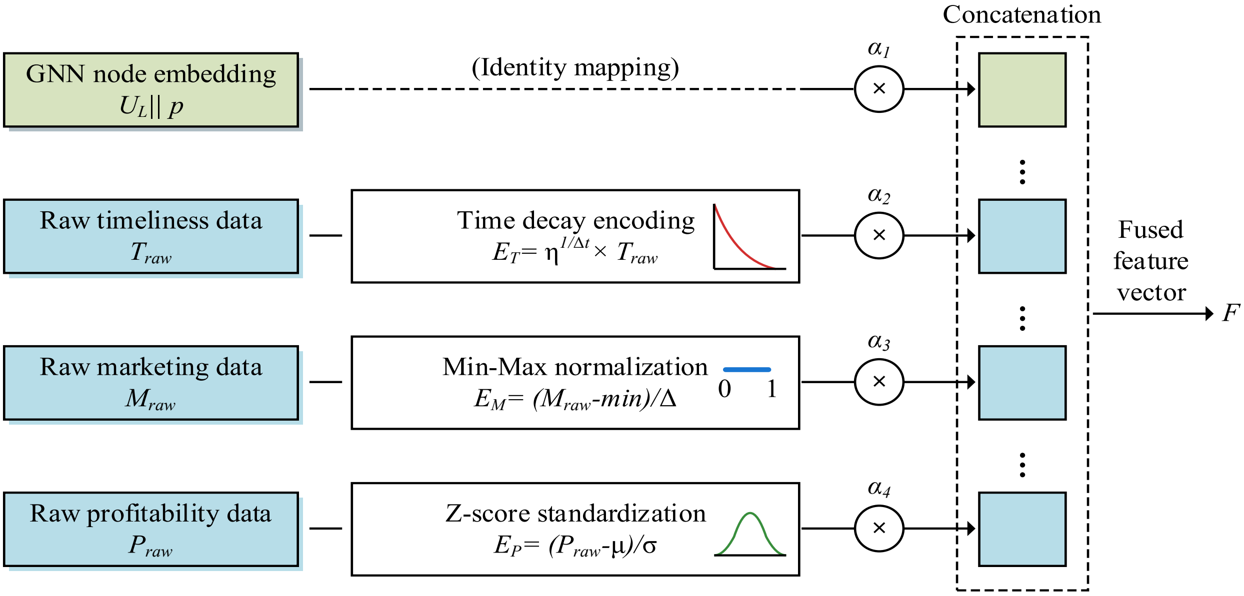


Fig. 4. Schematic illustration of the DFF mechanism

$$E_M = \frac{M_{raw} - M_{min}}{M_{max} - M_{min}} \tag{8}$$

In Eq. (8), E_M is the encoded marketing feature vector, and M_{raw} is the original feature value. M_{min} and M_{max} signify the minimum and maximum values of the marketing feature across the entire dataset, respectively. This normalization eliminates scale discrepancies between different promotional activities. For profitability features, which often follow a specific statistical distribution, Z-score standardization is utilized. The profitability feature encoding is shown in Eq. (9).

$$E_P = \frac{P_{raw} - \mu}{\sigma} \tag{9}$$

In Eq. (9), E_P represents the encoded profitability feature vector. P_{raw} is the original profitability value (e.g., profit margin). μ and σ denote the mean and standard deviation of the profit rate across all commodities in the dataset. This standardization ensures that profitability features follow a standard normal distribution (mean of 0, standard deviation of 1), making their weights comparable to other features during the fusion process. Finally, the encoded vectors are weighted and fused. As illustrated in the right section of Fig. 4, the model applies learnable attention weights to each feature stream before aggregating them via a concatenation operation. The overall DFF is shown in Eq. (10).

$$F = \mathrm{Concat}(\alpha_1(U_L \parallel p), \alpha_2 E_T, \alpha_3 E_M, \alpha_4 E_P) \tag{10}$$

In Eq. (10), F is the final fused feature vector. The terms α_1,α_2,α_3, and α_4 are learnable scalar weights (satisfying $\sum \alpha_i = 1$) that are automatically optimized during training. These weights quantify the relative importance of the GNN embedding, timeliness, marketing, and profitability features for the final prediction tasks. The function $\mathrm{Concat}(\cdot)$ represents the vector concatenation operation. The resulting vector F integrates the customer's long-term interests with short-term dynamic contexts and external business factors, providing a comprehensive input for the subsequent MTL module defined in Fig. 5. To collaboratively optimize multiple business goals, an MTL framework is designed, as illustrated in Fig. 5.

As shown in Fig. 5, the framework operates in a "Bottom-Up" manner. It utilizes the fused feature vector F generated by the DFF module (Eq. 10) as the Shared Feature Representation. This shared layer captures the underlying correlations between user preferences, churn risks, and value potential (Zhang et al., 2025). Based on this shared representation, the model branches into three task-specific "towers." Each tower consists of a specific Dense Layer (with learnable weights W and bias b) and an activation function tailored to the prediction target. The Preference Prediction (Main Task) aims to predict the probability of a user purchasing a product. It employs a Sigmoid activation function to map the output to the (0,1) interval. The formulation is shown in Eq. (11).

$$P_{pref} = \sigma(W_1 \cdot F + b_1) \tag{11}$$

In Eq. (11), P_{pref} denotes the predicted preference probability. W_1 and b_1 signify the learnable weight matrix and bias vector for the preference task, respectively. $\sigma(\cdot)$ signifies the Sigmoid activation function. The Churn Prediction (Auxiliary Task 1) identifies users at risk of leaving. Similar to the preference task, it uses a $Sigmoid$ activation to output a churn probability P_{churn}, as defined in Eq. (12).

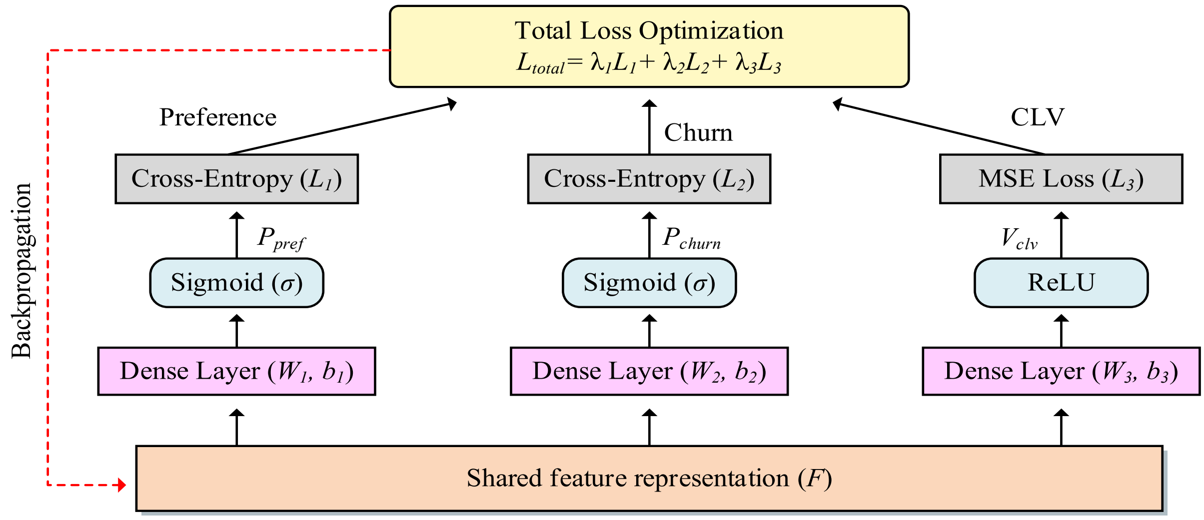


Fig. 5. Architecture of the MTL framework for collaborative optimization

$$P_{churn} = \sigma(W_2 \cdot F + b_2) \tag{12}$$

In Eq. (12), P_{churn} represents the predicted churn probability. W_2 and b_2 are the task-specific parameters. A value closer to 1 indicates a higher risk of customer churn within the observation window. The CLV Prediction (Auxiliary Task 2) estimates the potential financial value of the customer. Since CLV is a non-negative continuous variable, a Rectified Linear Unit ($ReLU$) activation function is adopted (Zhou et al., 2025). The prediction is shown in Eq. (13).

$$V_{clv} = ReLU(W_3 \cdot F + b_3) \tag{13}$$

In Eq. (13), V_{clv} is the predicted customer lifetime value. W_3 and b_3 are the parameters for the regression task. To ensure precise model training, the labels for these tasks are rigorously defined. For the churn prediction task, a user is labeled as "churned" (binary label 1) if no interaction is detected for 30 consecutive days. Otherwise, the label is 0. For the CLV estimation task, the label is defined as the normalized log-transformed sum of the user's transaction value over the subsequent 3-month period. The model is trained end-to-end by optimizing a Total Loss Function, which is a weighted sum of the individual task losses (Yie et al., 2024). As shown in the top block of Fig. 5, the total loss is shown in Eq. (14).

$$L_{total} = \lambda_1 L_1 + \lambda_2 L_2 + \lambda_3 L_3 \tag{14}$$

In Eq. (14), Where L_1 and L_2 are Cross-Entropy losses for the classification tasks (Preference and Churn), and L_3 is the Mean Squared Error (MSE) loss for the CLV regression task. λ_1,λ_2 and λ_3 are hyperparameters balancing the contribution of each task. As indicated by the red dashed line in Fig. 5, the error is minimized via Back-propagation, updating both the task-specific parameters (W_k,b_k) and the shared feature representation simultaneously. This total loss guides the optimization of model parameters through the back-propagation algorithm, ensuring that all tasks can learn simultaneously and reinforce one another, thereby improving recommendation accuracy while effectively addressing problems such as customer churn warning and value assessment. The advantage of this framework is that it integrates multiple related CRM tasks into a single model, avoiding the complexity of training models separately for each task, and achieving a balance between prediction accuracy and business goals. After completing the structural design of the MTL framework, the research still needs to integrate it into a complete training and optimization process. Therefore, the research will clarify the end-to-end training process of the model, comprehensively presenting the complete closed loop of data preparation, feature fusion, multi-task model training, and optimization through the total loss function, as shown in Fig. 6.

In Fig. 6, the training flowchart proposed in the study centers on intelligent model training and optimization, presenting a complete closed loop from data to commercial applications. The process first integrates multi-dimensional data via multi-source data collection and knowledge graph-building modules, providing a basis for model training. These data then enter the core dual-layer GNN model and feature fusion module. The dual-layer attention mechanism captures customers long-term and short-term preferences in layers and integrates dynamic features to enhance prediction accuracy (Sheng et al., 2025). The final model simultaneously predicts customer preferences, churn risk and CLV through MTL and prediction output modules, and recirculates the effect data through business feedback and closed-loop feedback to continuously optimize the model, forming a dynamic and adaptive intelligent system (Kim and Yoo, 2025). This complete process not only solves the data sparsity and model business adaptability, but also provides a solid foundation for the subsequent in-depth coupling of business goals and model output. The research also takes into account the balance between recommendation accuracy and corporate profits. The specific mechanism is shown in Fig. 7.

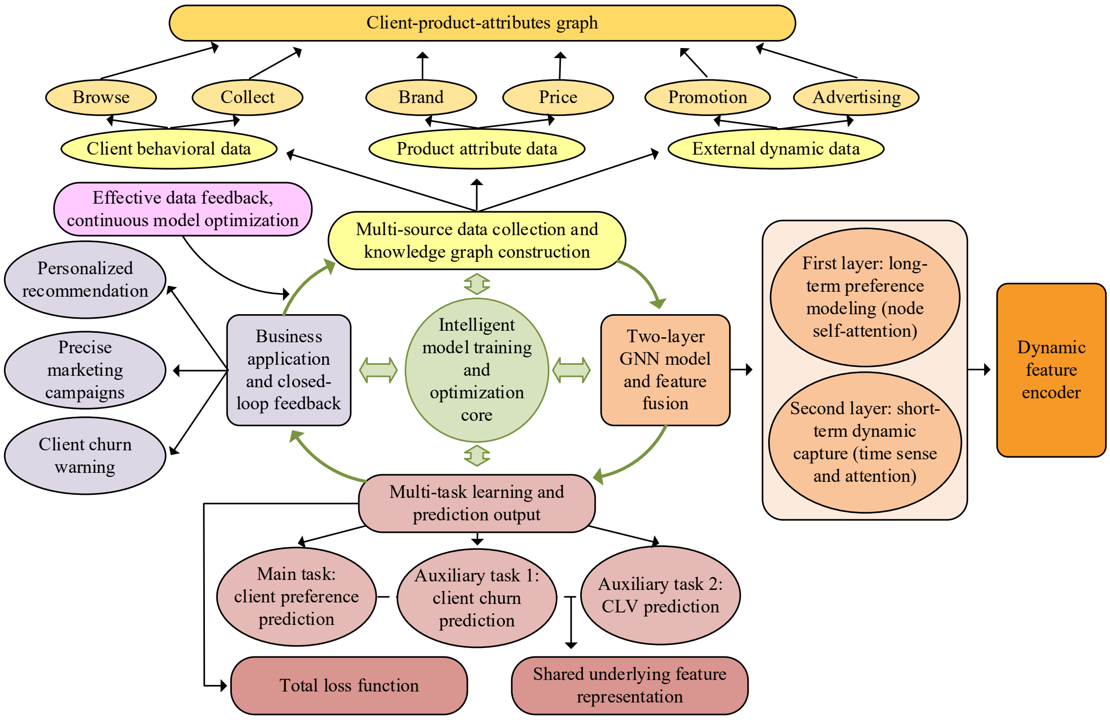


Fig. 6. Training flow chart based on DFF and MTL

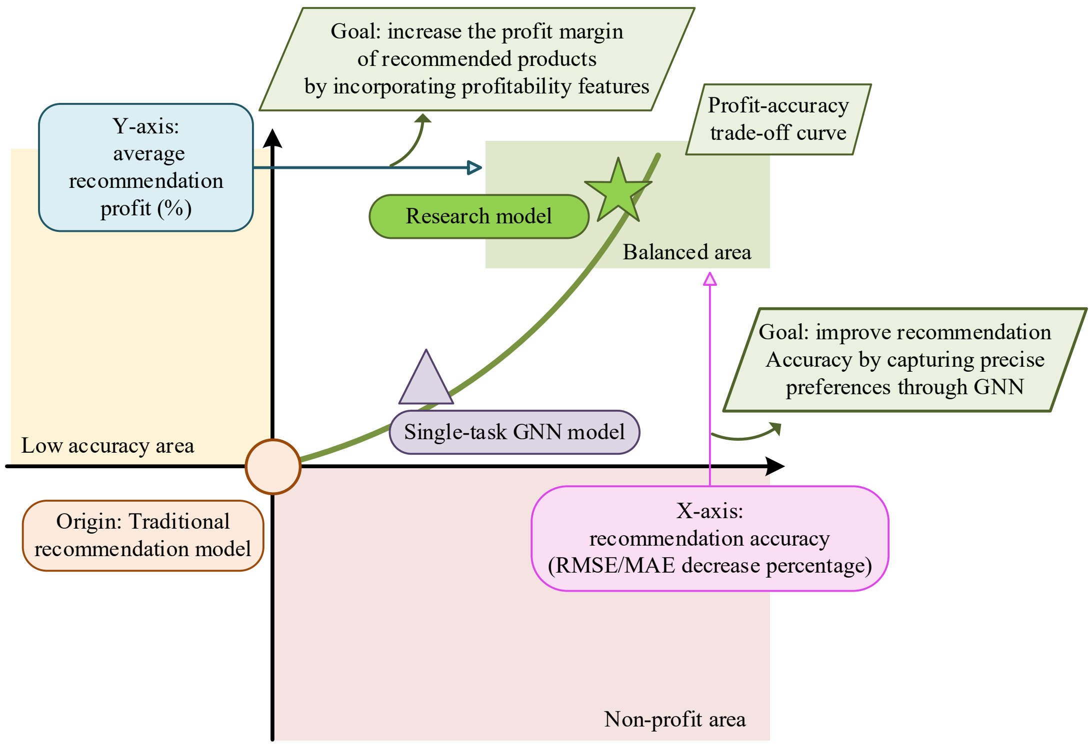


Fig. 7. Balance mechanism between profit and recommendation accuracy

As shown in Fig. 7, this mechanism takes balance as the core to quantify the comprehensive performance in a two-dimensional coordinate system. The X-axis represents recommendation accuracy, which reflects customer satisfaction. The Y-axis represents average recommendation profit, reflecting corporate profitability. By introducing the DFF module, including profitability features, the model can not only accurately predict customer preferences but also take product profit margins into account (Li et al., 2024a). In Fig. 7, the research model performance point is located in the upper right corner of the coordinate system, which is significantly better than the traditional model that only focuses on accuracy. The research model can effectively avoid the low-profit, low-accuracy area and the high-profit, low-accuracy area, thereby finding the best balance between the two (Huang et al., 2025). This mechanism successfully integrates customer experience and business value into a unified optimization objective, ensuring that the model can not only improve customer satisfaction but also deliver substantial economic returns to the enterprise in practical applications. The study named the proposed model as E-commerce Personalized Marketing Demand Prediction and Customer Relationship Management Model Based on Dual-Layer Attention Mechanism GNN, DFF and MTL (DLAGNN).

3. Results

3.1. Performance Evaluation and Analysis of DLAGNN Model

To verify the superior performance of the DLAGNN built on data mining technology, the study selected three representative models as comparison benchmarks, namely: GNN-Based Social Recommendation Model (GNN-SR), Time-Aware Recurrent Neural Network (TA-RNN), and Sequential Classification Model (SCM). The experiment's core data source is the Alibaba Tianchi dataset, which contains rich user behavior and product interaction data, providing a solid foundation for verifying the model's generalization ability. All experiments were conducted on a high-performance workstation equipped with an NVIDIA RTX 3090 GPU, 128GB of DDR4 RAM and a unified software environment built with Python 3.10 and the PyTorch 1.12.0 deep learning framework. In the data preprocessing stage, the study cleaned and anonymized the data set and constructed a customer-commodity-attribute relationship graph according to the model requirements to effectively capture the complex relationships between entities. Specifically, the data set consists of 100,000 user interaction logs and was randomly partitioned into training, validation, and test sets in an 8:1:1 ratio. This split ensures that the model is evaluated on unseen data, thereby guaranteeing the reliability of the generalization performance reported below. The study first evaluated the training efficiency and convergence performance, as presented in Fig. 8.

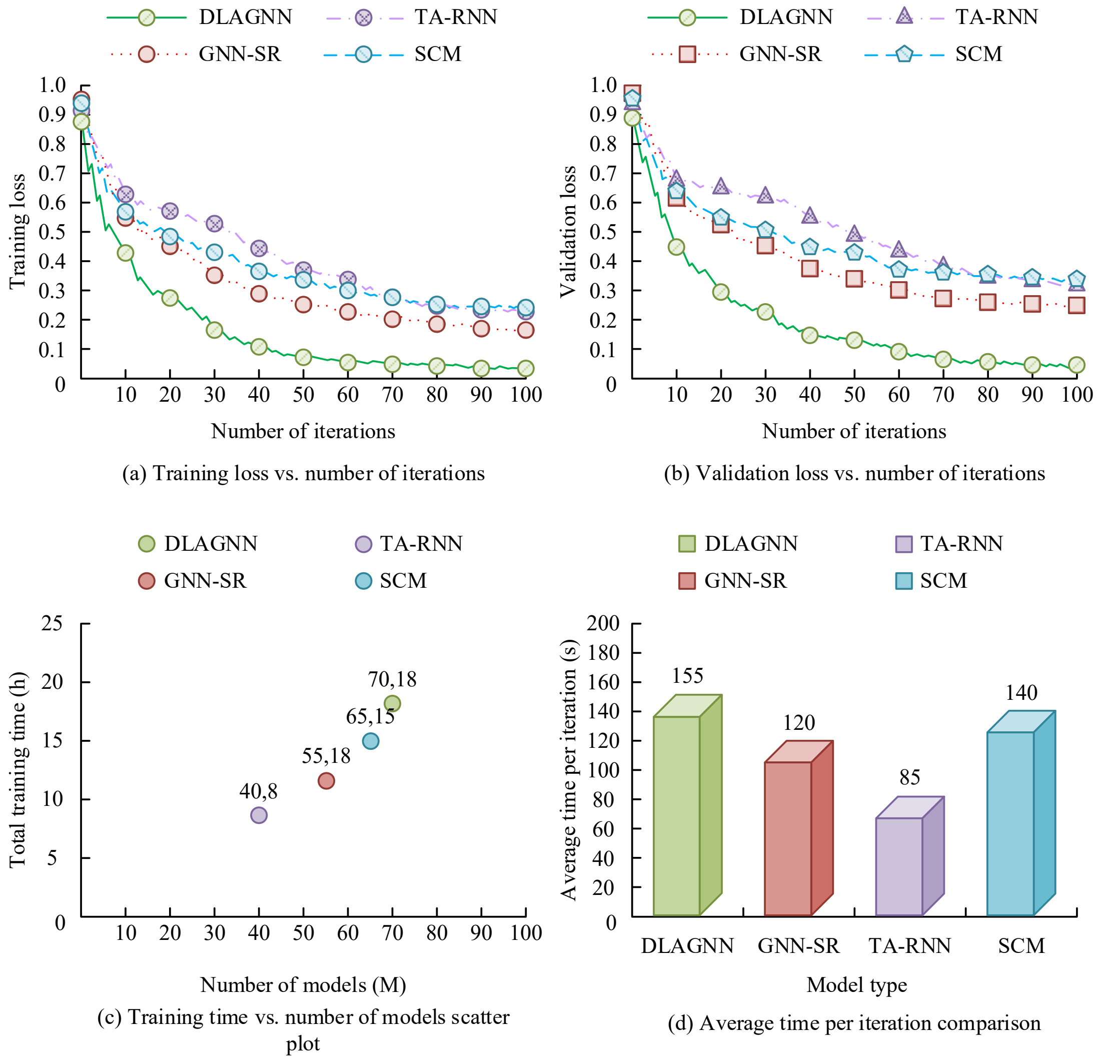


Fig. 8. Comparison of convergence rate and training efficiency of each model

As shown in Figs. 8(a) and 8(b), during the training and verification process, the loss value curve of the DLAGNN model dropped significantly faster than that of the comparison model, and basically converged after about 40 iterations. The loss value finally dropped to about 0.03, which was significantly lower than the 0.17-0.33 of other models. This shows that the DLAGNN model achieves higher optimization efficiency and can learn deeper features of the data faster. As shown in Fig. 8(c), the DLAGNN model had a parameter amount of 70M due to the dual-layer attention mechanism and MTL, which was slightly higher than that of other models. The total training time was 18 hours and did not grow linearly with its parameter amount, which reflected its efficient training strategy and convergence speed. In Fig. 8(d), the average single iteration time of the DLAGNN model was 155 seconds. Although it was about 70 seconds longer than that of the TA-RNN model with a smaller number of parameters, the time cost was completely reasonable considering its significant advantages in convergence speed and final performance. Taken together, the DLAGNN model shows stronger training efficiency and faster convergence speed while maintaining controllable computing costs. The research continues to comprehensively evaluate the accuracy in predicting personalized marketing demand, as presented in Fig. 9.

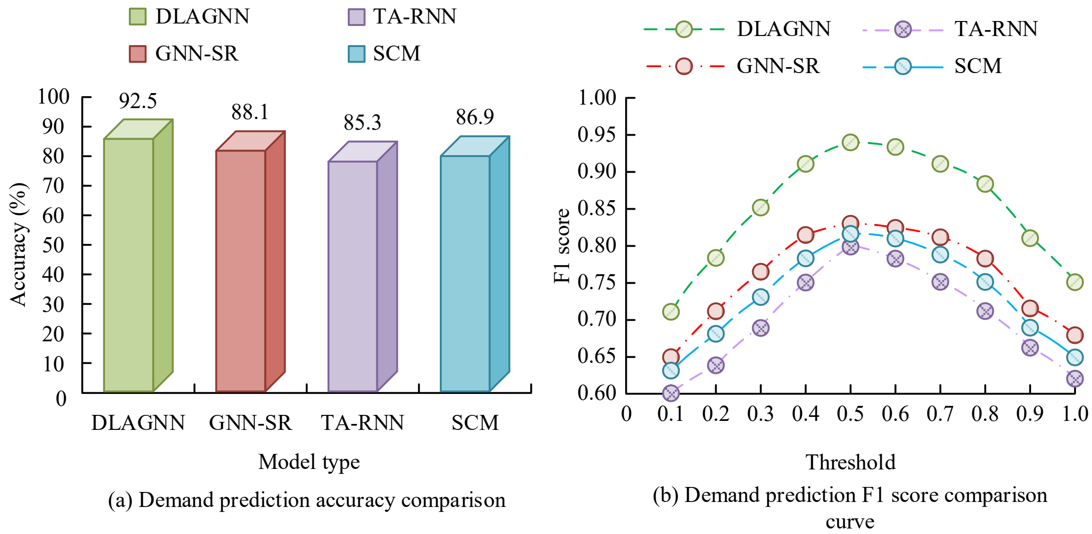


Fig. 9. Comparison of demand prediction accuracy of various models

As shown in Fig. 9(a), the DLAGNN model achieved excellent demand prediction accuracy, reaching up to 92.5%, significantly better than all comparison models. The accuracy of the GNN-SR model was 88.1%, TA-RNN was 85.3%, and SCM was 86.9%. This shows that the dual-layer attention mechanism and DFF strategy adopted by DLAGNN can more accurately capture customers' long- and short-term preferences, thereby making more accurate predictions. As shown in Fig. 9(b), the DLAGNN model maintained the highest F1 score under different prediction thresholds, especially reaching a peak of 0.94 near the threshold of 0.5, showing its best balance in prediction precision and recall. In contrast, the peak values of the F1 score curves of other models were lower and fluctuated greatly. The peak value of GNN-SR was about 0.86, while that of TA-RNN and SCM was only 0.80 and 0.82, respectively. This result strongly demonstrates the DLAGNN's ability to effectively address data sparsity and dynamic drift in customer needs, and its comprehensive predictive performance far exceeds that of the existing benchmark model. To verify the practical value of the DLAGNN in core CRM tasks, the study followed the Alibaba Tianchi data set and used its rich user behavior sequences to construct churn prediction task labels to verify the model's ability to identify potential churn risks, as presented in Fig. 10.

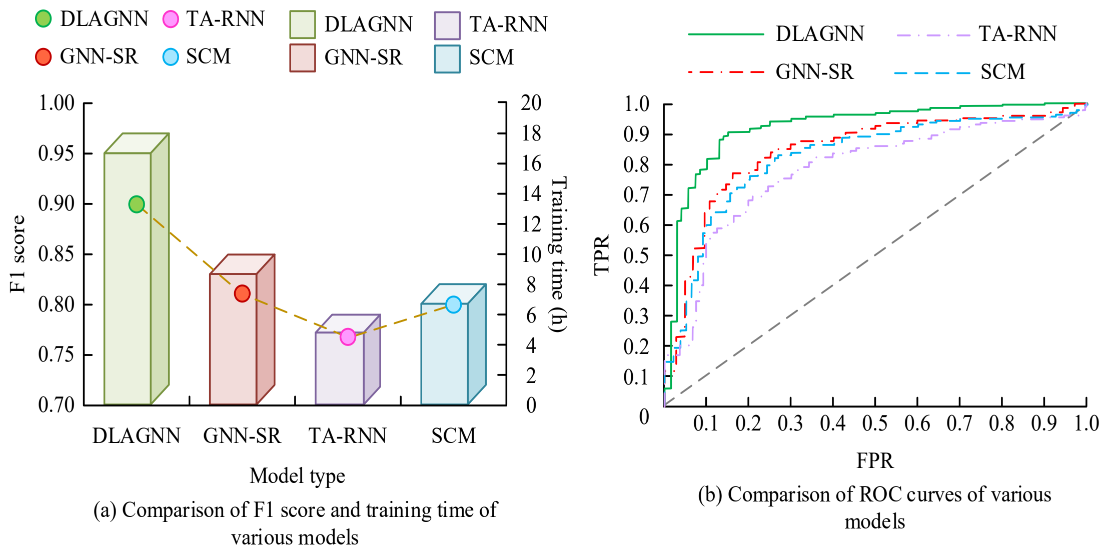
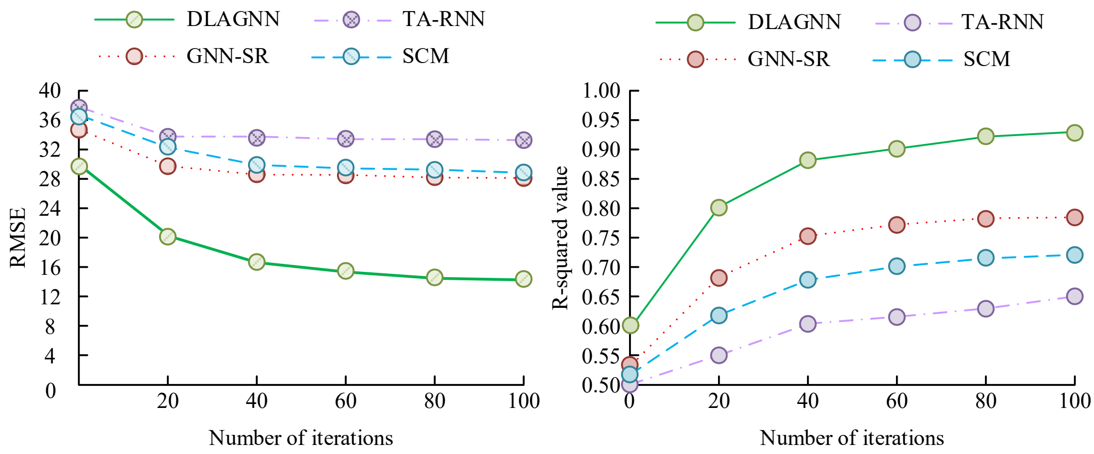


Fig. 10. Comparison of customer churn warning performance of each model

As shown in Fig. 10(a), in the customer churn warning task, the F1 score of the DLAGNN model was as high as 0.90, which was significantly better than that of GNN-SR (0.81), SCM (0.80) and TA-RNN (0.77). This shows that DLAGNN reached the best level in balancing the recognition accuracy and recall rate of lost customers. This also intuitively reflects the trade-off between performance and computational cost. Although the training time of DLAGNN was 18 hours, which was longer than that of other models, its improvement in prediction accuracy was significant. In Fig. 10(b), the ROC curve of the DLAGNN was significantly closer to the upper-left corner, and its Area Under the Curve (AUC) was the largest, indicating that the model had the strongest discriminative power across different classification thresholds. It can effectively distinguish between customer groups that are about to be lost and those that are retained, and provide enterprises with more accurate churn warnings, thereby improving the efficiency and effectiveness of CRM and enabling timely intervention measures to reduce customer churn. The study further constructed CLV evaluation task labels to verify each model's ability to predict customers' long-term profitability. The specific test results are shown in Fig. 11.



(a) Comparison of CLV prediction errors (b)Comparison of CLV prediction accuracy

Fig. 11. Comparison of CLV evaluation accuracy of each model

As shown in Fig. 11(a), in the CLV evaluation task, the Root Mean Square Error (RMSE) of the DLAGNN model decreased rapidly with the increase of iteration rounds, and converged to around 15.65 after about 80 iteration rounds, which was significantly lower than that of GNN-SR (28.15), TA-RNN (33.55), and SCM (29.05). This shows that the DLAGNN model can predict customers' long-term business value more accurately, with smaller errors and a closer match to the true value. As shown in Fig. 11(b), when evaluating the accuracy of CLV prediction, the R-squared value of the DLAGNN model rose rapidly with iteration rounds and stabilized at around 0.92 after about 80 iteration rounds, far exceeding that of GNN-SR (0.78), TA-RNN (0.64), and SCM (0.71). The closer the R-squared value is to 1, the stronger the model's ability to explain the dependent variable. This result demonstrates the superiority of the DLAGNN model in capturing key characteristics that affect CLV and in predicting long-term customer profitability. It provides a reliable basis for companies to formulate accurate customer stratification and resource allocation strategies, and further highlights its practical value in improving corporate business returns. To comprehensively assess the model's performance in real-world scenarios, specific business-oriented metrics are introduced alongside standard technical indicators. Mean Relative Absolute Error (MRAE) measures the average magnitude of errors in CLV predictions relative to actual values, reflecting the precision of financial estimates. Lead Accuracy Time Efficiency (LATE) quantifies the timeliness of churn warnings, defined as the average number of days between the predicted high-risk signal and the actual churn event. Value Conversion Improvement Percentage (VCIP) assesses the potential revenue uplift by simulating the change in conversion rate when marketing resources are allocated based on the model's recommendations. The study verifies the prediction accuracy, churn-warning capability, and customer-value assessment accuracy of each model, providing a quantitative basis for evaluating practicality in multi-task scenarios, as presented in Table 1.

In Table 1, for Marketing Resource Allocation Efficiency (MRAE), the DLAGNN reached 3.25, which was much higher than the highest value of 1.98 among the other models. This strongly demonstrates that DLAGNN, based on MTL and DFF strategies, can more effectively translate predictive insights into tangible business benefits and allocate marketing resources more precisely. In terms of Customer Churn Early Warning Time (LATE), DLAGNN could predict customer churn risks 24.6 days in advance on average, providing enterprises with a valuable retention window of about nine days longer than the GNN-SR model, thereby greatly improving CRM efficiency. In addition, in terms of Valuable Customer Identification Precision (VCIP), the DLAGNN model was far ahead, achieving an accuracy of 88.7%, indicating that its high reliability in CLV prediction tasks can help companies more accurately identify and target high-value customer groups. Although the number of parameters and training time of the DLAGNN are relatively high, its significant performance advantages across core business indicators demonstrate its strong value and effectiveness in practical applications.

3.2. Simulation Application Effect and Analysis of DLAGNN Model

To verify the effectiveness of the DLAGNN in practical applications, the study used the open source Taobao user behavior data set of Alibaba Tianchi platform to design three sets of simulation experiments, focusing on the three core business application scenarios of personalized product recommendation, customer churn intervention and CLV improvement. The study first focused on the key link in personalized product recommendation and conducted a marketing campaign effect simulation experiment. The study designed two experimental groups to compare the effects: The control group used the traditional marketing strategy based on the best-selling list to recommend the highest-selling products to customers. The experimental group used the trained DLAGNN model to generate personalized product recommendations based on each customer's predicted preferences. By comparing the performance of the two groups on key indicators such as purchase conversion rate, average customer price and MRAE, and the actual value of the DLAGNN model in improving marketing effects is quantified. The specific experimental results are shown in Fig. 12.

Table 1. Comparison of comprehensive performance of each model

Evaluation Dimension	Evaluation Metric	Unit	DLAGNN	GNN-SR	TA-RNN	SCM
Personalized Marketing	Marketing Resource Allocation Efficiency (MRAE)	/	3.25	1.98	1.57	1.83
	Model Training Duration	Hours	18	15	11	14
	Model Parameters	Million	70	55	25	42
	Prediction Latency	Sec/1,000 Predictions	12.3	9.8	8.5	10.1
Customer Relationship Management	Customer Churn Early Warning Time (LATE)	Days	24.6	15.2	11.8	14.7
	Valuable Customer Identification Precision (VCIP)	%	88.7%	75.1%	68.3%	72.9%
	Model Interpretability Score	1-10	7.5	6.2	4.8	5.5
	Data Integration Complexity	1-10	8.0	6.5	5.0	5.8
	Real-time Prediction Capability	1-10	8.9	7.6	6.8	7.1

As shown in Fig. 12(a), the recommendation conversion rate of the DLAGNN model increased rapidly with increasing marketing investment. Under the three investment levels of low, medium and high, the conversion rates of DLAGNN were 2.5%, 3.8% and 4.2%, respectively, which were much higher than the 0.9%, 1.5% and 1.8% of the control group. This shows that DLAGNN can more accurately match customer needs and achieve higher returns on marketing investment. As shown in Fig. 12(b), the customer transaction price distribution recommended by the DLAGNN model was generally higher than that of the control group. The median unit price for DLAGNN was US$23.26, whereas that of the control group was only US$16.25. This is because the DLAGNN model incorporates profitability features into its predictions, which can better balance customer preferences and product profitability and guide customers to purchase high-value products. As shown in Fig. 12(c), DLAGNN showed higher stability and performance in MRAE. The MRAE performance distribution of DLAGNN was concentrated between 3.1 and 3.5, which was much higher than that of the control group (0.9 to 1.3). This strongly demonstrates that the DLAGNN model can accurately allocate marketing resources, achieve higher returns with fewer resources, and deliver considerable commercial value to enterprises. The study further used the DLAGNN model to predict the probability of customer churn in real time and to design personalized intervention strategies, as presented in Table 2.

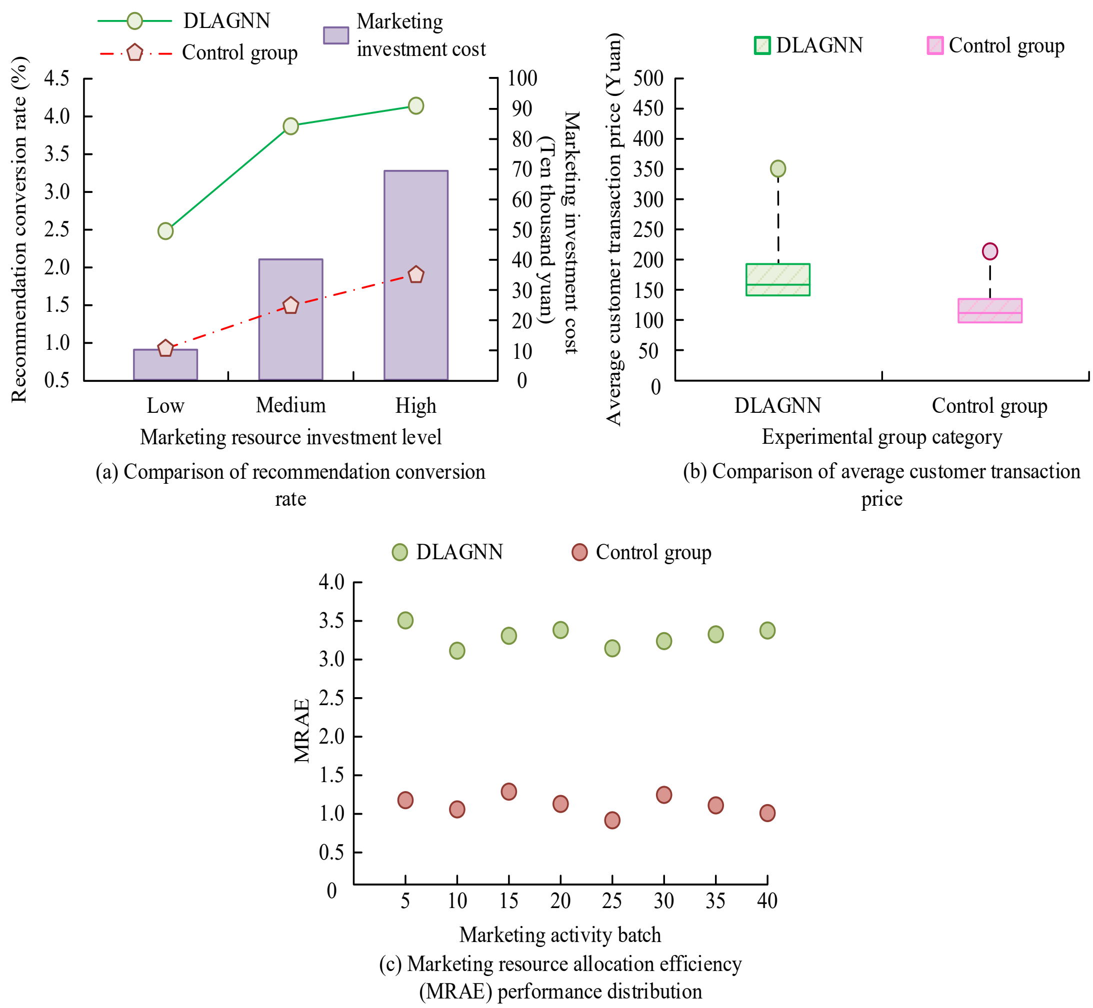


Fig. 12. Comparison of personalized marketing campaign effects predicted by DLAGNN model

In Table 2, the DLAGNN model demonstrated excellent practicality for early warning and intervention in customer churn. The model accurately identified 1,215 potential lost customers, and up to 97.12% of its predictions were consistent with the actual situation, demonstrating its extremely high accuracy. In addition, the DLAGNN could issue warnings an average of 24.6 days in advance, providing companies with sufficient time to implement effective retention strategies. The experimental results also confirmed the commercial value of the intervention. The intervention success rate of 22.8% brought in new revenue of US$12,713.43, and the Retention Cost (ROI) on investment was as high as 4.2. The DLAGNN also maintained a low false alarm rate (2.9%), effectively avoiding resource waste for non-churning customers and achieving a balance between customer experience and business efficiency. This proves that the DLAGNN is not only superior in predictive performance but also has significant economic value in actual commercial applications. To verify the applicability of the DLAGNN model to customer long-term value management, based on users' historical consumption behavior in the Alibaba Tianchi dataset, CLV tags are calculated and constructed, and the DLAGNN model is used to predict customers' potential CLV, and differentiated marketing strategies are designed accordingly. Finally, the changes in customer behavior before and after implementing the CLV strategy are compared to quantify the actual impact of the model on improving CLV, as presented in Table 3.

In Table 3, the DLAGNN model, through accurate prediction and stratification, significantly improved CLV by implementing differentiated marketing strategies for customer groups with different values. Low-value customers had the highest CLV improvement rate, at 27.04%, indicating that the model can effectively tap and activate the potential value of long-tail customers. The CLV of high-value and medium-value customers also increased by 21.74% and 21.26%, respectively, confirming the effectiveness of this strategy in consolidating existing value customers. In addition, the application of the DLAGNN led to significant improvements in macro business indicators, including an 18.5% decrease in the customer churn rate, a 15.2% increase in the repurchase rate, and a 0.6% increase in the proportion of high-value customers. Taken together, the ROI of this strategy was as high as 5.6, proving the great practical value and economic benefits of the DLAGNN model in guiding CLV management.

Table 2. Customer churn intervention simulation results of DLAGNN model

Alerting Metric	Predicted Value	Evaluation Result	Business Insight
Predicted Churning Customers	1,215	1,180 actual churns	The model identifies the vast majority of impending churns with 97.12% accuracy.
Early Warning Days	24.6 Days	Retention window increases by 68% compared to traditional methods.	Provides valuable time for the business to intervene, boosting retention success.
Retention Intervention Success Rate	22.8%	Out of 277 customers who received the intervention, they made a purchase within the next 30 days.	Personalized interventions are effective and the churn prediction is highly accurate.
New Revenue from Retained Customers	US$12,713.43	Retained customers generated additional revenue within the next 30 days.	The retention strategy yielded a direct economic return.
Reduced Customer Lifetime Value Loss	35.1%	Successfully retained customers reduced the expected CLV loss.	Proves the model's effectiveness in mitigating long-term business value loss.
Retention Cost (ROI)	4.2	For every US$1 invested in retention, US$4.2 in revenue was generated.	The intervention strategy has a high return on investment and is commercially viable.
False Positive Rate (FPR)	2.9%	Only 35 customers were incorrectly flagged as churning.	Low FPR avoids wasting resources on non-churning customers.

4. Discussion

The DLAGNN model, based on data mining technology, has achieved remarkable results in e-commerce personalized marketing and CRM. Its core advantages are the dual-layer attention mechanism and the DFF strategy, which effectively address data sparsity, dynamic drift in customer preferences, and the disconnect between business goals and technical indicators.

First, the DLAGNN model performs well at customer demand prediction, thanks to its unique dual-layer attention-based GNN architecture. The first-layer attention mechanism effectively captures and models customers long-term preferences through weighted aggregation of customer historical interaction records. This method avoids the problem that traditional collaborative filtering or rule-based grouping methods are difficult to capture individual subtle preferences and data sparsity. Just like the previous research by Zhao et al. (2022), which used a sequence learning architecture to improve the accuracy of customer segmentation, the GNN architecture further deepens the understanding of complex interactive relationships by modeling customers, products and attributes as graphs. The second-layer attention mechanism focuses on capturing short-term dynamic demand, modeling recent behavioral sequences through time-aware attention, and incorporating contextual information such as marketing activities to accurately predict demand drift. This design allows the model not only to identify customers stable long-term interests but also to respond quickly to instantaneous changes in demand. This is similar to the idea of Lin et al. (2024), who used the enhanced conformer model to process dynamic data in online ride-hailing demand prediction. The rapid convergence and high accuracy show that this hierarchical and dynamic modeling method can efficiently extract valuable patterns from massive data, thereby significantly improving the accuracy and effectiveness of personalized recommendations.

Second, the excellent performance of the DLAGNN in CRM tasks confirms the great value of the MTL and DFF strategies. Integrating customer preference prediction, churn warning, and CLV prediction into a unified framework enables the model to simultaneously optimize multiple business goals, which is in sharp contrast to the traditional cumbersome way of building models independently for each task. The DFF module combines key business factors such as timeliness, external

Table 3. Simulation results of customer stratification and CLV improvement strategy in DLAGNN model

Customer Segment	Avg. CLV Before Strategy	Avg. CLV After Strategy	CLV Increase Ratio	Strategy Description	Avg. Purchase Frequency (per month)	Avg. Transaction Value
High-Value	US$223.35	US$271.92	21.74%	Exclusive new product recommendations, VIP member benefits	3.5 times	US$77.69
Medium-Value	US$99.99	US$121.26	21.26%	Targeted discount coupons, personalized product recommendations	2.1 times	US$57.65
Low-Value	US$18.37	US$23.33	27.04%	First-purchase discounts, limited-time promotions	1.2 times	US$16.99
Overall Customers	US$113.92	US$138.83	21.86%	Differentiated marketing strategies	2.3 times	US$50.77
High-Value Customer Ratio	15.2%	15.8%	+0.6%	Growth in the high-value customer base	/	/
Churn Rate Reduction	/	18.5%	/	Churn rate significantly decreased after strategy implementation	/	/
Repurchase Rate Increase	/	15.2%	/	Repurchase intention strengthened across different customer segments	/	/
ROI	/	5.6	/	The strategy yielded a high return	/	/

marketing and profitability with the underlying feature representation of the model, ensuring that the prediction results are not only technically accurate, but also commercially feasible. For example, in the churn warning task, the DLAGNN model can predict customer churn risk 24.6 days in advance on average, which is about 9 days more precious retention time than that of the GNN-SR model. This result far exceeds traditional customer churn early warning models and highlights the ability of the DLAGNN model to provide a priori insights at key decision points. In addition, the model's high accuracy in CLV assessment (R-square value as high as 0.92) proves that it can provide enterprises with reliable customer stratification basis to guide differentiated resource allocation and marketing strategies. This finding echoes the research by Ferrer-Estévez et al. (2023), who sorted out knowledge in the field of CRM and emphasized its importance in sustainable development. By taking profitability characteristics into account, the DLAGNN model successfully found the optimal balance between recommendation accuracy and corporate profits, thereby allocating marketing resources precisely and efficiently.

Despite the remarkable results of the study, there are still some limitations. First, the DLAGNN model is highly dependent on multi-source heterogeneous data and requires high-quality, multi-dimensional data inputs, which may be difficult to obtain in some enterprises with imperfect data infrastructure. Second, although the model's complex architecture and large number of parameters have improved performance, they have also increased computational cost and training time, which may limit its applicability in scenarios with extremely high real-time requirements or limited computing resources. Future research can explore lightweight model architectures, such as those based on knowledge distillation or pruning techniques, to reduce computational overhead while maintaining performance. In addition, the research primarily focuses on customers' online behavioral data, but their offline behaviors, social media interactions, and other external macroeconomic factors may also influence their consumption decisions. Future research can further explore multi-modal data fusion and integrate more data sources to build a more comprehensive customer portrait. Meanwhile, drawing on research on the application of GNNs in social recommendation systems by Sharma et al. (2024), the model accounts for users' social relationships and community influence to better understand customer preferences and behaviors. This will help build a more robust and generalized intelligent marketing framework, thereby better empowering the long-term development of e-commerce companies.

5. Conclusion

As e-commerce transforms into a customer-centric model, traditional CRM is unable to meet the needs of precision marketing. To solve the limitations of traditional methods in e-commerce personalized marketing demand prediction and CRM, the research took data mining technology as the core to construct a DLAGNN model that integrated dual-layer attention mechanisms, GNN, DFF and MTL. The research represented the interactive relationship by constructing a customer-commodity-attribute knowledge graph, capturing customers' long- and short-term preferences with dual-layer attention, and incorporating timeliness and profitability features. The three major tasks of preference prediction, churn warning and CLV prediction were simultaneously optimized. Performance verification showed that the demand prediction accuracy of DLAGNN was 92.5%, the F1 score for churn warning was 0.90, the R2 value for CLV prediction was 0.92, and the MRAE was 3.25, all of which were significantly better than those of the comparison models. In the simulation test, its personalized recommendation achieved a maximum conversion rate of 4.2%, a churn intervention return on investment of 4.2%, and an overall increase in customer CLV of 21.86%, demonstrating its practical value. In summary, the research has demonstrated that the DLAGNN model can effectively address data sparsity and preference drift, achieving a deep coupling between technical accuracy and business goals. It is important to note that the reported improvements in ROI and business impact are derived from simulation experiments based on historical data. While these results demonstrate the model's theoretical potential, actual economic gains in a production environment may vary with real-time market dynamics and execution costs. However, the research still has limitations, such as high dependence on data quality and high computational cost. Follow-up research can explore the model's lightweight architecture, integrate multimodal data and users social relationships, and further improve its generalization ability and practicality.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

Declaration of Artificial Intelligence (AI) Tools

The author used ChatGPT solely for language editing and to improve readability. The author reviewed and verified all content and takes full responsibility for the accuracy and integrity of the manuscript.References

Alves Gomes, M. and Meisen, T. (2023). A review on customer segmentation methods for personalized customer targeting in e-commerce use cases. *Inf. Syst. E-Bus. Manag,* 21(3), 527-570. https://doi.org/10.1007/s10257-023-00640-4

Berahmand, K., Bahadori, S., Abadeh, M. N., Li, Y., and Xu, Y. (2024). SDAC-DA: Semi-supervised deep attributed clustering using dual autoencoder. *IEEE transactions on knowledge and data engineering*, 36(11), 6989-7002. https://doi.org/10.1109/TKDE.2024.3389049

Chen, S., Zhang, Y., and Yang, Q. (2024). Multi-task learning in natural language processing: An overview. *ACM Computing Surveys*, 56(12), 1-32. https://doi.org/10.1145/3663363

Ferrer-Estévez, M. and Chalmeta, R. (2023). Sustainable customer relationship management. *Marketing Intelligence & Planning*, 41(2), 244-262. https://doi.org/10.1108/MIP-06-2022-0266

Gao, C., Zheng, Y., Li, N., Li, Y., Qin, Y., Piao, J., and Li, Y. (2023). A survey of graph neural networks for recommender systems: Challenges, methods, and directions. *ACM Transactions on Recommender Systems*, 1(1), 1-51. https://doi.org/10.1145/3568022

Groumpos, P. P. (2023). Artificial intelligence: Issues, challenges, opportunities and threats. In *Conference on Creativity in Intelligent Technologies and Data Science*, 19-33. Cham: Springer International Publishing. https://doi.org/10.47852/bonviewAIA3202689

Hou, B., Cai, Z., Wu, K., Yang, T., and Zhou, T. (2023). 6scan: A high-efficiency dynamic internet-wide ipv6 scanner with regional encoding. *IEEE/ACM Transactions on Networking,* 31(4), 1870-1885. https://doi.org/10.1109/TNET.2023.3233953

Huang, L., Huang, X., Zou, H., Gao, Y., Wang, C., and Yu, P. S. (2025). Knowledge-Reinforced Cross-Domain Recommendation. *IEEE Transactions on Neural Networks and Learning Systems*, 36(7), 12880-12894. https://doi.org/10.1109/TNNLS.2024.3500096

Kim, H. J. and Yoo, S. J. (2025). Disturbance observer-based adaptive chainlike filter approach for prescribed-time consensus tracking of nonlinear multiagent systems via dynamic state and input triggering. *IEEE Transactions on Cybernetics*. https://doi.org/10.1109/TCYB.2025.3576352

Li, Q., Wang, Y., and Lv, F. (2024a). Semantic correlation attention-based multiorder multiscale feature fusion network for human motion prediction. *IEEE transactions on cybernetics*, 54(2), 825-838.https://doi.org/10.1109/TCYB.2022.3184977

Li, X., Yan, H., Cui, K., Li, Z., Liu, R., Lu, G., and Hon, C. (2024b). A novel hybrid YOLO approach for precise paper defect detection with a dual-layer template and an attention mechanism. *IEEE Sensors Journal*, 24(7), 11651-11669. https://doi.org/10.1109/JSEN.2024.3356356

Lin, H., He, Y., Liu, Y., Gao, K., and Qu, X. (2024). Deep demand prediction: An enhanced conformer model with cold-start adaptation for origin–destination ride-hailing demand prediction. *IEEE Intelligent Transportation Systems Magazine*, 16(3), pp.111-124. https://doi.org/10.1109/MITS.2023.3309653

Mao, J., Xie, J., Hu, Z., Deng, L., Wu, H., and Hao, Y. (2023). Sustainable development through green innovation and resource allocation in cities: Evidence from machine learning. *Sustainable Development*, 31(4), 2386-2401. https://doi.org/10.1002/sd.2516

Mykhalchenko, H. and Tytarenko, M. (2023). Data analytics and personalized marketing strategies in E-commerce platforms. *Futurity Economics & Law,* 3 (3), 114–138 [online]. https://doi.org/10.57125/FEL.2023.09.25.07

Rahmani, S., Baghbani, A., Bouguila, N., and Patterson, Z. (2023). Graph neural networks for intelligent transportation systems: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 24(8), 8846-8885. https://doi.org/10.1109/TITS.2023.3257759

Raji, M. A., Olodo, H. B., Oke, T. T., Addy, W. A., Ofodile, O. C., and Oyewole, A. T. (2024E-commerce and consumer behavior: A review of AI-powered personalization and market trends. *GSC advanced research and reviews*, 18(3), 066-077. https://doi.org/10.30574/gscarr.2024.18.3.0090

Sharma, K., Lee, Y. C., Nambi, S., Salian, A., Shah, S., Kim, S. W., and Kumar, S. (2024). . A survey of graph neural networks for social recommender systems. *ACM Computing Surveys*, 56(10), 1-34. https://doi.org/10.1145/3661821

Sheng, Z., Cao, Y., Yang, Y., Feng, Z., Shi, K., Huang, T., and Wen, S. (2025Residual temporal convolutional network with dual attention mechanism for multilead-time interpretable runoff forecasting. *IEEE Transactions on Neural Networks and Learning Systems*, 36(5), 8757-8771. https://doi.org/10.1109/TNNLS.2024.3411166

Wang, D., Li, T., Deng, P., Zhang, F., Huang, W., Zhang, P., and Liu, J. (2023). A generalized deep learning clustering algorithm based on non-negative matrix factorization. *ACM Transactions on Knowledge Discovery from Data*, 17(7), 1-20 .https://doi.org/10.1145/3584862

Wang, Q., Wang, S., Zhuang, D., Koutsopoulos, H., and Zhao, J. (2024aUncertainty quantification of spatiotemporal travel demand with probabilistic graph neural networks. *IEEE Transactions on Intelligent Transportation Systems*, 25(8), 8770-8781. https://doi.org/10.1109/TITS.2024.3367779

Wang, Y., Li, S., Liu, C., Wang, K., Yuan, X., Yang, C., and Gui, W. (2024b). Multiscale feature fusion and semi-supervised temporal-spatial learning for performance monitoring in the flotation industrial process. *IEEE Transactions on Cybernetics*, 54(2), 974-987. https://doi.org/10.1109/TCYB.2023.3295852

Xu, Y., Lyu, Y., Xiong, G., Wang, S., Wu, W., Cui, H., and Luo, J. (2023). Adaptive feature fusion networks for origin-destination passenger flow prediction in metro systems. *IEEE Transactions on Intelligent Transportation Systems*, 24(5), 5296-5312. https://doi.org/10.1109/TITS.2023.3239101

Yao, M., Zhao, G., Zhang, H., Hu, Y., Deng, L., Tian, Y., and Li, G. (2023a). Attention spiking neural networks. *IEEE transactions on pattern analysis and machine intelligence*, 45(8), 9393-9410. https://doi.org/10.1109/TPAMI.2023.3241201

Yao, Z., Xia, S., Li, Y., and Wu, G. (2023b). Cooperative task offloading and service caching for digital twin edge networks: A graph attention multi-agent reinforcement learning approach. *IEEE Journal on Selected Areas in Communications*, 41(11), 3401-3413. https://doi.org/10.1109/JSAC.2023.3310080

Yie, S. Y., Kim, K. M., Bae, S., and Lee, J. S. (2024Effects of loss functions and supervision methods on total-body PET denoising. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 8(4), 379-390. https://doi.org/10.1109/TRPMS.2023.3334276

Zhang, Y., Wu, X., Fang, Q., Qian, S., and Xu, C. (2023). Knowledge-enhanced attributed multi-task learning for medicine recommendation. *ACM Transactions on Information Systems*, 41(1), 1-24. https://doi.org/10.1145/3527662

Zhang, Z., Wang, T., Hu, Z., Yang, L., and Li, H. (2025). DEMENTIA: a hybrid attention-based multimodal and multi-task learning framework with expert knowledge for Alzheimer's disease assessment from speech. *IEEE Journal of Biomedical and Health Informatics,* 29(4), 2957-2968. https://doi.org/10.1109/JBHI.2024.3509620

Zhao, L., Zuo, Y., and Yada, K. (2022). Sequential classification of customer behavior based on sequence-to-sequence learning with gated-attention neural networks. *Advances in Data Analysis and Classification,* 17(3), 549-581. https://doi.org/10.1007/s11634-022-00517-3

Zhong, L., Wu, J., Li, Q., Peng, H., and Wu, X. (2023). A comprehensive survey on automatic knowledge graph construction. *ACM Computing Surveys*, 56(4), 1-62. https://doi.org/10.1145/3618295

Zhou, M., Bai, Y., Zhang, W., Yao, T., and Zhao, T. (2025). Interactive conversational head generation. *IEEE Transactions on Pattern Analysis and Machine Intelligence.,* 47(8), 6673-6686. https://doi.org/10.1109/TPAMI.2025.3562651

Junyi Yang attained two master's degrees, a Master of Information Technology and Master of Business Administration from James Cook University, Australia, in 2024. Currently, he holds the position of full-time faculty member at the School of Artificial Intelligence, Chongqing Jianzhu College, China. His core research focuses on data mining, computer vision, and the application of artificial intelligence and digital marketing. He is also deeply engaged in software engineering and has participated in multiple research and industry-collaborative projects in this area.