Artificial Intelligence and Computing Innovations

Intelligence Science Publishing

https://www.ispublishing.com

ISSN: 3070-4944

Vo.4, No.4, 2024

Intelligent Decision-Making for Advertising Recommendation via Data Mining and Graph Neural Networks

Zichen Pan[1]*

[1]Columbia University, New York, USA

**Corresponding author: peter18nj@gmail.com*

Abstract: This paper addresses the challenges of data sparsity, complex interactions, difficulty in fully characterizing latent preferences, and lack of structural awareness in the decision-making process of advertising recommendation tasks. It proposes an intelligent decision-making method for advertising recommendation that combines hard data mining and graph neural networks. First, this method extracts multi-source information, including user, ad, and contextual information, from the original ad interaction logs. Then, it uses a hard data mining mechanism to identify high-value behavioral patterns and key relationship clues to reduce noise interference and enhance the ability to express effective information. Based on this, a heterogeneous interaction graph for advertising recommendation scenarios is constructed, unifying user interests, ad attributes, and environmental information into a graph structure modeling framework. Graph propagation learning is used to mine high-order dependencies between nodes, thereby more accurately representing potential associations in complex recommendation scenarios. Simultaneously, a graph reweighting and adaptive fusion mechanism is introduced to strengthen key structural information and improve the synergy between initial features and deep graph representations, making the resulting decision representation more consistent with the actual needs of advertising recommendation tasks. Furthermore, based on a unified decision modeling strategy, the matching relationship between users and ads is effectively estimated, thereby completing ad ranking and recommendation output. Comparative experimental results show that the proposed method achieves superior performance on multiple recommendation evaluation indicators, verifying its effectiveness and practicality in intelligent decision-making tasks for advertising recommendation.

Keywords: Heterogeneous interaction modeling; high-order relation learning; ad ranking optimization; decision representation fusion

1. Introduction

Against the backdrop of the deepening digital economy and the continuous evolution of platform business models, advertising recommendation has become a crucial link connecting user demand, merchant supply, and platform revenue. Compared to traditional uniform delivery methods, intelligent advertising recommendation for complex scenarios emphasizes a comprehensive portrayal of changes in user interests, semantic relationships of content, evolution of behavioral trajectories, and differences in contextual environments, thereby achieving more accurate supply-demand matching and more efficient resource allocation [1, 2]. The scale of data faced by advertising systems continues to expand, and the data sources are becoming increasingly diverse, including explicit behavioral information such as clicks, browsing, dwell time, and conversions, as well as implicit structural features such as session paths, time dependencies, group associations, and contextual feedback. In this context, how to extract stable, effective, and decision-valuable information from massive, heterogeneous, and noisy data has become one of the core issues in the intelligent development of advertising recommendation [3, 4].

To address these issues, the combination of hard data mining and graph neural networks provides a new research approach for intelligent decision-making in advertising recommendation [5]. Hard data mining emphasizes further identifying high-value patterns, key association rules, potential group structures, and deep behavioral preferences from complex data, helping to overcome the limitations of surface statistical correlation and improve the ability to identify users' true needs and advertising value characteristics [6]. Graph neural networks, on the other hand, can model high-order dependencies between nodes in a graph structure space, and through neighborhood aggregation, relation propagation, and structural representation learning, more comprehensively reflect the interaction topology characteristics in the advertising ecosystem. The integration of these two approaches not only enhances the recommendation model's ability to express complex relationships but also provides more targeted and systematic support for decision-making processes such as ad placement, ranking optimization, traffic allocation, and strategy control. Table 1 summarizes the main practical drivers and theoretical value of this research direction, demonstrating that it possesses both practical necessity and room for methodological innovation.

Table 1. Overview of the Research Value of Intelligent Decision-Making in Advertising Recommendation Combining Deep Data Mining and Graph Neural Networks

Research Dimension	Main Content	Practical Significance
Data Value Mining	Identify deep preferences and key patterns from high-dimensional, sparse, and heterogeneous behavioral data.	Improve user demand understanding and ad matching accuracy.
Relational Structure Modeling	Construct multi-level relational networks among users, ads, content, and scenarios.	Strengthen the representation and propagation modeling of complex interaction relationships.
Intelligent Decision Optimization	Apply learned representations to ad delivery, ranking, bidding, and resource allocation.	Enhance platform revenue, advertising effectiveness, and user experience.
Methodological Integration Innovation	Integrate deep mining mechanisms with graph learning mechanisms to build a novel recommendation framework.	Promote advertising recommendations from correlation matching toward structured intelligent decision-making.

Based on the development trend of intelligent recommendation systems, this research focuses on intelligent decision-making in advertising recommendations that combines hard data mining with graph neural networks, possessing significant theoretical and practical value. From a theoretical perspective, this direction helps expand the unified modeling framework for deep data patterns and complex relationship structures in advertising recommendations, promoting the evolution of recommendation methods from feature-driven to a synergistic evolution of relationship-driven and decision-driven approaches [7]. From an application perspective, this direction can better adapt to the practical needs of advertising platforms in areas such as precise targeting, dynamic optimization, performance improvement, and risk control, providing support for building a high-quality, efficient, and highly adaptable advertising recommendation system. Especially in an environment of continuously growing data scale, rapidly changing user needs, and intensifying market competition, research in this direction has important practical significance for improving the intelligence level of advertising recommendation systems, promoting collaborative optimization of the platform ecosystem, and releasing the value of data elements.

2. Datasets and Dataset Preprocessing

2.1 Dataset

This paper selects the publicly available Criteo Display Advertising Challenge dataset as the data foundation for research on intelligent decision-making in advertising recommendation. This dataset, designed for display ad click-through rate prediction, is derived from real-world ad traffic scenarios and contains a week's worth of sample records. Each sample corresponds to one ad display event and provides a binary label indicating whether the user clicked the ad. In terms of feature composition, the dataset includes 13 numerical features and 26 categorical features, comprehensively reflecting user, ad exposure, and contextual interaction information. It exhibits typical characteristics such as high dimensionality, sparsity, and heterogeneity, thus demonstrating a high degree of compatibility with the topic of intelligent decision-making in advertising recommendation that combines hard data mining and graph neural networks. This dataset has been publicly released on the Kaggle platform and has long been used in research related to CTR prediction, ad ranking optimization, and recommendation modeling, demonstrating good openness, representativeness, and academic reproducibility.

2.2 Dataset Preprocessing

This paper further presents the comparison results before and after data preprocessing, as shown in Figure 1.

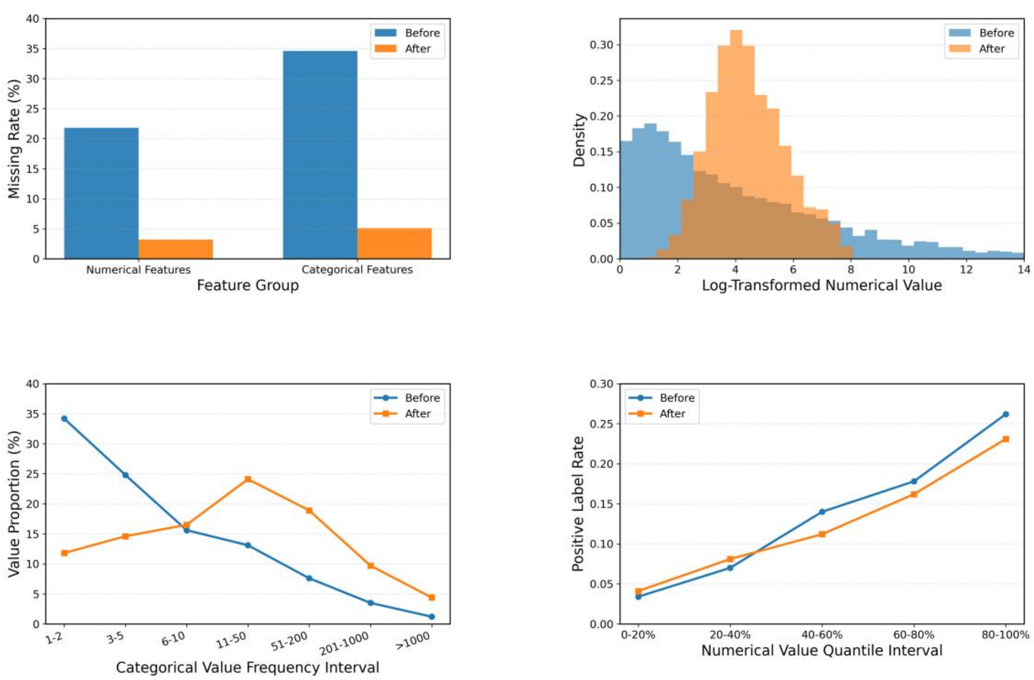

Figure 1. Results of data preprocessing

Figure 1 illustrates the statistical results before and after data preprocessing from four aspects: changes in the missing rate, distribution pattern of numerical features, frequency structure of class values, and stratification trend of positive sample rate. Specifically, the upper left figure reflects the overall changes in the missing rate of numerical and class features. After preprocessing, the missing rates of both types of features decreased significantly, indicating that missing value imputation and outlier sample cleaning effectively improved data integrity. The upper right figure shows a comparison of the distribution of numerical features after transformation. It can be seen that the data distribution before preprocessing was significantly skewed, while the distribution after preprocessing is more concentrated and smooth, which helps to reduce the interference of extreme values and improve the learnability of features. The lower left figure shows the changes in the proportion of class value frequency intervals. After preprocessing, the proportion of low-frequency classes is compressed, and the structure of mid-frequency classes is clearer, indicating that the sparse class distribution is effectively regularized, which helps to reduce the modeling noise caused by class sparsity. The lower right figure depicts the trend of positive sample rate changes in different numerical quantile intervals. After preprocessing, the overall curve is smoother and more monotonic, indicating that the correspondence between sample labels and features is more stable. Overall, after preprocessing, the data showed significant improvements in completeness, distribution stability, clarity of category structure, and consistency of tag associations, laying a more reliable data foundation for the training and representation learning of the subsequent intelligent decision-making model for advertising recommendation.

3. Methodology

The proposed method is designed to support intelligent decision-making in advertising recommendation by jointly exploiting deep data mining and graph neural representation learning. Rather than treating impression records as isolated samples, the whole recommendation environment is organized as a heterogeneous interaction graph in which users, advertisements, categories, and contextual states are connected through behavior-driven relations, because ad exposure, click response, and preference evolution are inherently dependent on structured interactions rather than independent observations. From this perspective, the original high-dimensional and sparse advertising logs are first transformed into a relational graph to preserve cross-entity dependency, while the accompanying numerical and categorical attributes are embedded into a unified feature space to enhance semantic consistency across different data sources. This paper further presents the overall model architecture, as shown in Figure 2.

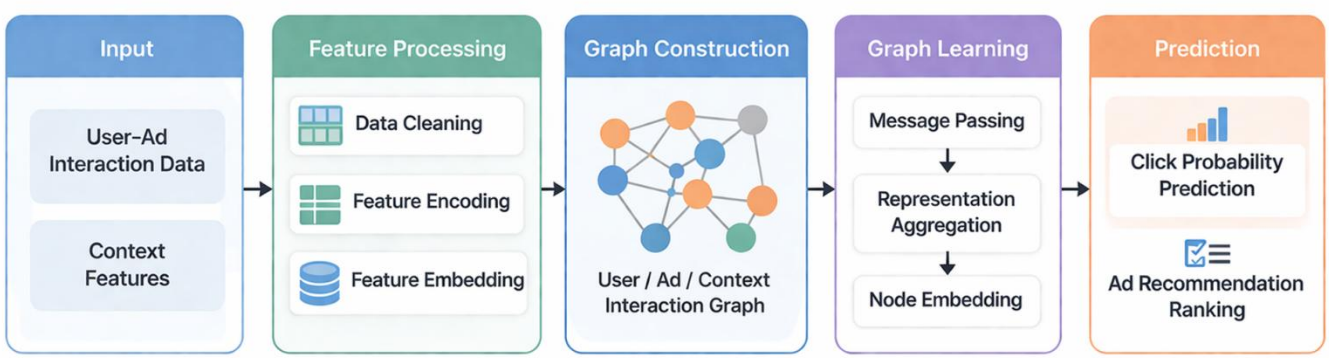

Figure 2. Overall model architecture

Let the constructed advertising graph be denoted as:

$$\mathcal{G} = (\mathcal{V}, \mathcal{E}, \boldsymbol{X}, \boldsymbol{T}) \tag{1}$$

where $\mathcal{V}$ denotes the set of graph nodes, $\mathcal{E}$ represents the set of interaction edges, $\boldsymbol{X}$ is the node attribute matrix, and $\boldsymbol{T}$ records edge types or behavioral semantics associated with different recommendation events. To avoid directly feeding noisy and weakly informative raw records into the graph encoder, a hard data mining module is introduced before graph propagation so that critical behavioral patterns, stable preference signals, and high-value structural cues can be emphasized in advance. Accordingly, the initial node representation is obtained through a joint embedding transformation:

$$\boldsymbol{H}^{(0)} = \phi(\boldsymbol{X}\boldsymbol{W}_0 + \boldsymbol{b}_0) \tag{2}$$

in which $\boldsymbol{W}_0$ and $\boldsymbol{b}_0$ control the projection from heterogeneous raw attributes into a shared latent space and $\phi(\cdot)$ denotes a nonlinear activation used to strengthen expressive capacity. This design is meaningful because it allows the subsequent graph learning process to start from semantically aligned representations rather than fragmented raw features, thereby improving the quality of downstream decision signals in advertising recommendation.

Beyond feature transformation alone, the method further emphasizes the extraction of hard patterns hidden in large-scale advertising behavior data, since user feedback in real recommendation platforms is usually sparse, biased, and highly imbalanced. A mining score is therefore assigned to each node or interaction instance to estimate its structural importance and behavioral value, allowing the model to focus more on informative relations that carry stronger decision relevance. Formally, the mining strength associated with the node v_i is defined as:

$$s_i = \alpha d_i + \beta c_i + \gamma p_i \tag{3}$$

where d_i denotes the normalized interaction density of node v_i, c_i measures cross-context behavioral consistency, and p_i reflects preference concentration estimated from historical response records, while α, β, and γ balance the contributions of these three complementary factors. On this basis, the original edge connectivity is reweighted to suppress weak or noisy relations and to preserve more reliable recommendation dependencies, which yields:

$$\tilde{a}_{ij} = a_{ij}\left(1 + \lambda \frac{s_i + s_j}{2}\right) \tag{4}$$

where a_{ij} is the original adjacency value between nodes v_i and v_j and λ adjusts the influence of mined importance on graph connectivity. Such a reweighting mechanism is crucial because it enables the graph structure to reflect not only whether interactions occurred, but also how valuable these interactions are for advertising decision-making, thereby making the learned topology more suitable for recommendation-oriented inference.

Once the graph has been refined by hard data mining, a graph neural propagation process is employed to capture high-order dependency patterns among users, advertisements, and contextual entities. Instead of relying on shallow pairwise matching, the model aggregates multi-hop neighborhood information so that latent preference transfer, crowd-level influence, and content association can be encoded within node representations. The layer-wise propagation rule is written as:

$$\boldsymbol{H}^{(l+1)} = \sigma\left(\widehat{\boldsymbol{D}}^{-\frac{1}{2}}\widehat{\boldsymbol{A}}\widehat{\boldsymbol{D}}^{-\frac{1}{2}}\boldsymbol{H}^{(l)}\boldsymbol{W}^{(l)}\right) \tag{5}$$

where $\widehat{\boldsymbol{A}} = \widetilde{\boldsymbol{A}} + \boldsymbol{I}$ is the adjacency matrix with self-loops, $\widehat{\boldsymbol{D}}$ is its degree matrix, $\boldsymbol{W}^{(l)}$ is the trainable transformation at layer l, and $\sigma(\cdot)$ introduces nonlinear modeling ability during message passing. Since different neighbors do not contribute equally to recommendation decisions, an adaptive fusion strategy is adopted after propagation to combine the mined importance with structural representation learning, and the final decision-oriented embedding of the node v_i is expressed as:

$$\boldsymbol{z}_i = \eta_i \boldsymbol{h}_i^{(L)} + (1 - \eta_i)\boldsymbol{h}_i^{(0)} \tag{6}$$

where $\boldsymbol{h}_i^{(L)}$ is the representation after L graph layers, $\boldsymbol{h}_i^{(0)}$ is the initial embedded feature, and $\eta_i = \mathrm{sigmoid}(s_i)$ is an adaptive coefficient generated from the mining score. By combining propagated structural semantics with the original attribute signal in this manner, the representation becomes more robust to oversmoothing while retaining the most decision-relevant information discovered during the mining stage.

Finally, intelligent advertising recommendation is realized through a prediction layer that maps the learned embeddings into click-oriented decision probabilities, making the model suitable for ad ranking and recommendation optimization under complex interaction settings. Given a user node u and an advertisement node m, their interaction probability is estimated by:

$$\hat{y}_{um} = sigmoid(\boldsymbol{z}_u^{\top}\boldsymbol{Q}\boldsymbol{z}_m + b) \tag{7}$$

where $\boldsymbol{Q}$ is a trainable bilinear parameter matrix used to characterize cross-entity compatibility, and b is a global bias item for calibration. To ensure that the learned decision boundary remains consistent with the observed click behavior, the whole framework is optimized by the binary cross-entropy objective:

$$\mathcal{L} = -\frac{1}{N}\sum_{(u,m)\in\Omega}[y_{um}log\hat{y}_{um} + (1 - y_{um})log(1 - \hat{y}_{um})] \tag{8}$$

where Ω denotes the set of observed user-advertisement pairs, $y_{um} \in \{0,1\}$ is the ground-truth response label, and N is the number of training instances. Through this joint optimization process, the proposed method integrates data hard mining, graph structure refinement, high-order dependency aggregation, and probabilistic decision modeling into a unified framework, which is beneficial for uncovering deep interaction patterns in advertising data and for improving the rationality and adaptivity of recommendation decisions under complex commercial scenarios.

4. Experimental Results and Analysis

4.1 Experimental Setup

To ensure the stability and reproducibility of the model training and evaluation process, the experiments were conducted in a unified hardware and software environment, with specific configurations shown in Table 2.

Table 2. Experimental Settings of the Proposed Method

Category	Item	Setting
Hardware	GPU	NVIDIA GeForce RTX 4090
Hardware	CPU	Intel Xeon Silver 4314
Hardware	System Memory	128 GB
Software	Operating System	Ubuntu 22.04
Software	Programming Language	Python 3.10
Software	Deep Learning Framework	PyTorch 2.1.0
Software	Acceleration Toolkit	CUDA 12.1, cuDNN 8.9
Training Hyperparameter	Batch Size	1024
Training Hyperparameter	Learning Rate	0.001
Training Hyperparameter	Optimizer	Adam
Training Hyperparameter	Weight Decay	1e-5
Training Hyperparameter	Dropout Rate	0.3
Model Hyperparameter	Hidden Dimension	128
Model Hyperparameter	Graph Learning Layers	2
Training Hyperparameter	Epochs	100
Training Strategy	Early Stopping Patience	10
Training Strategy	Random Seed	42

On the hardware side, the training platform used an NVIDIA GeForce RTX 4090 GPU, an Intel Xeon Silver 4314 CPU, and 128 GB of system memory to meet the computational requirements of large-scale sample processing, graph structure construction, and iterative propagation of graph neural networks in the advertising recommendation task. On the software side, the experiments were implemented based on the Ubuntu 22.04 operating system, Python 3.10, and PyTorch 2.1.0, combined with CUDA 12.1 and cuDNN 8.9 to accelerate the model training process. In terms of parameter settings, the batch size was set to 1024, the learning rate to 0.001, the optimizer to Adam, the weight decay to 1e-5, the dropout rate to 0.3, the number of graph learning layers to 2, the hidden dimension to 128, and the number of training epochs to 100, to achieve a reasonable balance between training efficiency and representational power. Meanwhile, the patience of the early stopping strategy was set to 10, and the random seed was fixed at 42, thereby reducing the impact of training fluctuations on the experimental results and improving the consistency of the overall experimental process.

4.2 Experimental Results

To further evaluate the overall performance of the proposed method in intelligent decision-making scenarios for ad recommendation, representative works closely related to this paper's research direction are selected for comparative analysis. These methods cover related research areas such as click-through rate prediction, ad ranking modeling, feature interaction learning, and graph structure recommendation. Considering that recommendation tasks focus more on ranking quality and hit rate, Table 3 uses four classic recommendation metrics—HR, NDCG, MRR, and MAP—for horizontal comparison, and displays the performance differences between the proposed method and other related methods under the same evaluation dimension.

Table 3. Comparative Results of Different Methods

Method	HR	NDCG	MRR	MAP
Li et al. [8]	0.71	0.48	0.42	0.40
Ouyang et al. [9]	0.73	0.50	0.43	0.41
Wu et al. [10]	0.74	0.50	0.44	0.42
Shen et al. [11]	0.74	0.51	0.45	0.42
Zhai et al. [12]	0.75	0.52	0.46	0.43
Zhou et al. [13]	0.73	0.49	0.44	0.41
Wang et al. [14]	0.72	0.49	0.43	0.41
Ours	0.77	0.54	0.47	0.45

The overall results show that the proposed method outperforms other methods across various recommendation evaluation metrics, demonstrating its superior expressive power in modeling the complex relationships between users and advertisements and in uncovering latent interest structures. By integrating deep pattern mining and graph structure modeling mechanisms, this method effectively enhances the capture of high-order interaction information, resulting in more consistent recommendation results in terms of relevance and ranking rationality. Furthermore, the constructed representation maintains a stable advantage across different evaluation dimensions, further indicating that the method possesses good generalization ability and decision reliability in complex advertising recommendation scenarios, thus providing stronger technical support for intelligent advertising delivery.

To further analyze the impact of each key component on the overall intelligent decision-making performance of advertising recommendation, this paper conducts an ablation study focusing on the core aspects of the proposed method. The relevant results are shown in Table 4. Specifically, the ablation settings strictly correspond to the hard data mining mechanism, graph reweighting strategy, graph propagation learning process, and adaptive fusion mechanism in the method design, thus clearly revealing the actual roles of different modules in high-order interaction modeling, structural information transmission, and decision representation optimization.

Table 4. Ablation Study Results of the Proposed Method

Ablation Setting	HR	NDCG	MRR	MAP
w/o Hard Data Mining	0.73	0.50	0.44	0.42
w/o Graph Reweighting	0.74	0.51	0.45	0.43
w/o Graph Propagation Learning	0.72	0.49	0.43	0.41
w/o Adaptive Fusion	0.75	0.52	0.46	0.44
Ours	0.77	0.54	0.47	0.45

The experimental results demonstrate the strong effectiveness of the overall framework in intelligent decision-making tasks for advertising recommendation. This method effectively combines deep data pattern mining with graph structure relationship modeling, enabling a more comprehensive expression of the complex relationships between user interests, advertising attributes, and contextual information. Furthermore, the adaptive fusion mechanism enhances the consistency between the representation results and the decision objective. Overall, the proposed algorithm exhibits strong advantages in recommendation relevance, ranking quality, and result stability, indicating that this method can provide more reliable technical support for intelligent recommendation in complex advertising scenarios.

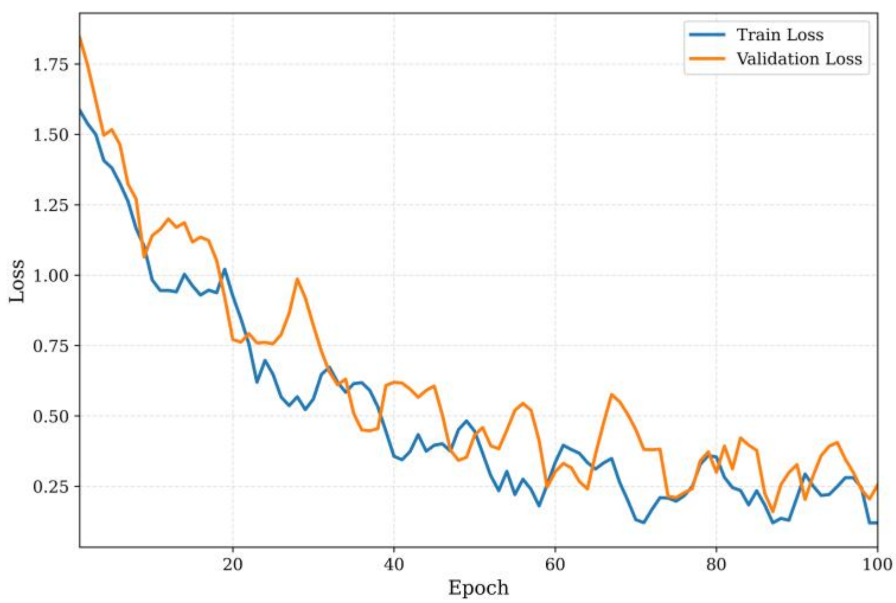

Figure 3. Experimental results of the loss function varying with epoch

As shown in Figure 3, the proposed algorithm maintains good optimization stability and convergence during training, with the loss value remaining within a reasonable range throughout the training phase, indicating that the model achieves effective coordination between feature representation and structural modeling. Furthermore, the method also demonstrates reliable generalization ability during the validation phase, suggesting that the constructed representation can adapt well to unseen data and reduce the risk of overfitting. Overall, the algorithm exhibits strong learning ability and stability in complex advertising recommendation tasks, providing a solid foundation for subsequent intelligent decision-making.

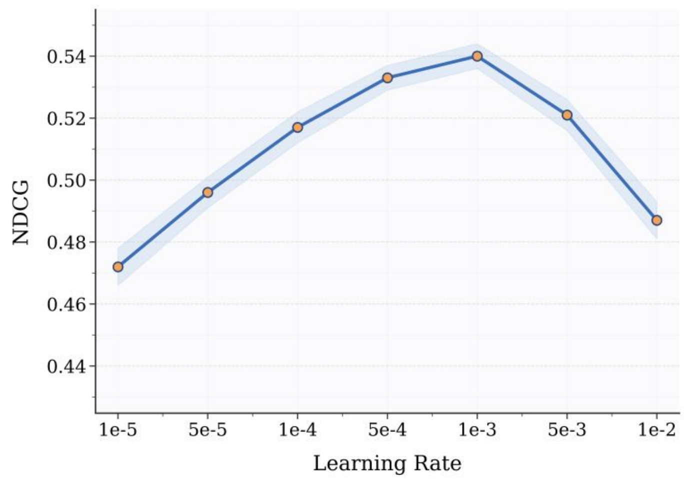

Figure 4. The impact of different learning rate settings on NDCG

The proposed algorithm maintains high recommendation ranking quality under different learning rate settings, demonstrating its good stability and adaptability during parameter updates, as shown in Figure 4. By effectively combining deep data pattern mining and graph structure representation learning, this method can continuously maintain high-quality feature representation during training, making the model robust in modeling complex interaction relationships. Furthermore, the constructed optimization mechanism helps improve the model's consistent performance under various training conditions, further demonstrating the algorithm's reliability and practical value in intelligent decision-making tasks for advertising recommendations.

5. Conclusion

This paper addresses the real-world challenges of high-dimensional sparsity, complex relationships, and long decision chains in advertising recommendation scenarios. It investigates an intelligent decision-making method for advertising recommendation that combines hard data mining and graph neural networks, aiming to improve the representational capabilities and decision quality of advertising recommendation systems in complex interactive environments. Addressing the difficulty of traditional methods in fully revealing the deep relationships between users, advertisements, and context, this paper adopts a unified design from two levels: deep data pattern mining and graph structure relationship modeling. It organically combines high-value behavioral information extraction, key relationship screening, interaction structure reconstruction, and higher-order dependency propagation to construct a unified decision-making framework for intelligent advertising recommendation. This method not only enhances the ability to characterize the latent semantic relationships and structural information in complex advertising ecosystems but also helps improve the overall performance of recommendation results in terms of relevance, ranking rationality, and decision consistency, thus providing new research ideas for refined modeling in advertising recommendation tasks.

From an overall research value perspective, the significance of this work lies not only in the structural integration and decision optimization at the methodological level but also in its effective response to the core needs of advertising recommendation application scenarios. As digital platform traffic organization methods continue to evolve, advertising recommendation systems have gradually evolved from simple click prediction tools into key intelligent infrastructures that balance user experience, merchant revenue, and platform resource allocation efficiency. In this context, recommendation methods capable of simultaneously understanding multi-source data relationships, mining deep behavioral patterns, and generating stable decision outputs have significant application potential. The framework proposed in this paper theoretically advances the evolution of advertising recommendation tasks from shallow feature matching to deep relationship reasoning and structured intelligent decision-making. In application, it provides a more supportive technical path for precise ad placement, ad ranking optimization, traffic value mining, commercial conversion improvement, and platform operation collaboration. This research has strong promotional value and practical significance for intelligent marketing, e-commerce platforms, content distribution systems, and related commercial recommendation scenarios.

Looking to the future, intelligent decision-making in advertising recommendation still has considerable room for expansion. On the one hand, with the continuous growth of user behavior sequences, cross-domain content semantics, and dynamic context environments, future research can further strengthen temporal dependency modeling, multimodal information fusion, and dynamic graph structure update mechanisms to enhance the model's ability to perceive real-time interest migration and scene changes. On the other hand, in real-world platform applications, advertising recommendation systems also need to balance interpretability, fairness, robustness, and deployment efficiency. Therefore, future research can focus on maintaining recommendation quality while further exploring lighter structural designs, more transparent decision-making methods, and optimization strategies more suitable for online inference. Simultaneously, extending this research framework to broader intelligent recommendation and business decision-making tasks is expected to have a more profound impact on related fields such as digital advertising, personalized marketing, content matching, and intelligent platform operation, and drive the continuous development of advertising recommendation technology towards higher quality, higher reliability, and higher application value.

References

[1] L. Sang, H. Li, Y. Zhang, et al., "AdaGIN: Adaptive Graph Interaction Network for Click-Through Rate Prediction", ACM Transactions on Information Systems, vol. 43, no. 1, pp. 1-31, 2024.

[2] Z. Xiao, L. Yang, T. Zhang, W. Jiang, W. Ning, Y. Yang, "Deep Evolutional Instant Interest Network for CTR Prediction in Trigger-Induced Recommendation," Proceedings of the 17th ACM International Conference on Web Search and Data Mining (WSDM), 2024.

[3] Y. Yang, C. Wen, H. Cui, Y. Sun and Y. Liu, "Learning Unified Multi-Granularity Representations for Backend Anomaly Detection and Causal Localization", 2024.

[4] Y. Yang, B. Chen, C. Zhu, et al., "AIE: Auction Information Enhanced Framework for CTR Prediction in Online Advertising", Proceedings of the 18th ACM Conference on Recommender Systems, pp. 633-642, 2024.

[5] J. Pan, W. Xue, X. Wang, et al., "Ads Recommendation in a Collapsed and Entangled World", Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp. 5566-5577, 2024.

[6] W. Guo, R. Tang, Y. Ye, Z. Li, X. He, "Dual Graph Enhanced Embedding Neural Network for CTR Prediction," arXiv preprint arXiv:2106.00314, 2021..

[7] H. Cui,"Task Graph Modeling and Constraint-Aware Scheduling for Reliable Large Language Model Agent Execution," 2024.

[8] Z. Li, Z. Cui, S. Wu, et al., "Fi-GNN: Modeling Feature Interactions via Graph Neural Networks for CTR Prediction", Proceedings of the 28th ACM International Conference on Information and Knowledge Management, pp. 539-548, 2019.

[9] W. Ouyang, X. Zhang, S. Ren, et al., "Learning Graph Meta Embeddings for Cold-Start Ads in Click-Through Rate Prediction", Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 1157-1166, 2021.

[10] K. Wu, W. Bian, Z. Chan, et al., "Adversarial Gradient Driven Exploration for Deep Click-Through Rate Prediction", Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp. 2050-2058, 2022.

[11] Q. Shen, H. Wen, W. Tao, et al., "Deep Interest Highlight Network for Click-Through Rate Prediction in Trigger-Induced Recommendation", Proceedings of the ACM Web Conference 2022, pp. 422-430, 2022.

[12] P. Zhai, Y. Yang and C. Zhang, "Causality-Based CTR Prediction Using Graph Neural Networks", Information Processing & Management, vol. 60, no. 1, Art. no. 103137, 2023.

[13] G. Zhou, X. Zhu, C. Song, et al., "Deep Interest Network for Click-Through Rate Prediction", Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 1059-1068, 2018.

[14] R. Wang, B. Fu, G. Fu, et al., "Deep & Cross Network for Ad Click Predictions", Proceedings of the ADKDD'17, pp. 1-7, 2017.