ISSN 2616-5902 Vol. 7, Issue 11: 134-140, DOI: 10.25236/AJBM.2025.071118

Data Management Innovation in the e-Commerce Field: Innovation in Personalized Recommendation Algorithms and User Experience Optimization

Li Mu

Changchun Humanities and Sciences College, Changchun, Jilin Province, 130117, China
254904998@qq.com

Abstract: *In the current context of information overload and increasingly diverse user needs on e-commerce platforms, collaborative filtering and content recommendation algorithms struggle to deeply analyze users' dynamic and fine-grained intent. To this end, this study proposes and implements a personalized recommendation framework that integrates dynamic knowledge graphs with causal inference. First, this paper constructs a dynamic product-attribute-scenario knowledge graph that integrates real-time user behavior sequences. Next, causal inference techniques are introduced, and finally, an exploration-exploitation strategy based on proximal policy optimization (PPO) is discussed. Experimental results show that during a four-week test, the framework achieved an average session duration of 43 minutes for the experimental group. In a bias correction experiment, it more accurately identified and recommended potentially high-quality long-tail products. In an exploration-exploitation balance experiment, the compound reward strategy achieved a 47.2% success rate in exploring new categories in the fourth training cycle. These results demonstrate that the proposed personalized recommendation framework can effectively enhance user stickiness and loyalty, significantly improve the diversity and fairness of the recommendation system, and demonstrate its superiority in discovering users' potential interests.*

Keywords: *E-commerce Data Management; Recommendation System; Causal Inference; Reinforcement Learning; Personalized Recommendation Algorithm*

1. Introduction

Against the backdrop of the rapid growth of e-commerce platforms, the amount of information users encounter continues to expand, and the supply of goods and services is highly diverse. However, this has also led to prominent issues such as homogeneous recommendation results, fixed user interests, and difficulty unlocking long-tail value. While traditional collaborative filtering and content recommendation methods have improved matching efficiency to a certain extent, they struggle to accurately capture dynamic user preferences and are unable to balance exploration and fairness. These challenges make ensuring accurate recommendations while stimulating user demand in an environment of information overload a key issue for the long-term and healthy development of e-commerce platforms.

To address these challenges, this paper proposes a personalized recommendation framework that integrates dynamic knowledge graphs, causal inference, and deep reinforcement learning. This approach leverages dynamic knowledge graphs to fine-grainedly model user interests, effectively eliminates data bias using causal inference techniques, and incorporates an intrinsic curiosity mechanism into reinforcement learning strategies to balance exploration and exploitation. This research not only provides new solutions for optimizing the diversity, fairness, and long-term value of recommendation systems, but also offers a viable technical path for core data management and user experience enhancement on e-commerce platforms.

The full paper is organized as follows: First, we describe the relevant research progress and theoretical foundations to lay the foundation for the framework proposed in this paper. Second, we introduce in detail the construction of dynamic knowledge graphs, the causal inference debiasing mechanism, and the reinforcement learning strategy that integrates intrinsic curiosity. Then, we design and implement three experiments to verify the effectiveness of the method in enhancing long-term user value, correcting biases, and balancing exploration and utilization. Finally, we summarize the research

ISSN 2616-5902 Vol. 7, Issue 11: 134-140, DOI: 10.25236/AJBM.2025.071118

conclusions and look forward to future development directions.

2. Related Work

To solve the data management and algorithm optimization problems in e-commerce recommendation systems, various methods have been proposed. For example, Zhang [1] designed an e-commerce personalized recommendation algorithm based on Hadoop to address the data sparsity, scalability and real-time problems in collaborative filtering recommendation algorithms. As the popularity of e-commerce brings business opportunities to enterprises, Gulzar et al. [2] tried to apply ordered clustering algorithms to alleviate the cold start and data sparsity problems in recommendation systems. Loukili et al. [3] developed a personalized recommendation method based on association rules and frequent pattern growth algorithms, which achieved good results. Shankar et al. [4] proposed an intelligent recommendation system based on ensemble learning, which plays a role in customer analysis in social media, online business and e-commerce. Wu and Chi [5] proposed a multi-dimensional recommendation system for e-commerce enterprises with comprehensive functions. Wei [6] took the e-commerce scenario as an example and designed and optimized an online recommendation system based on machine learning algorithms. Necula and Păvăloaia [7] explored the current status and future trends of artificial intelligence applications in e-commerce recommendation systems. With the popularization of information technology and mobile devices, Kim et al. [8] developed a variety of recommendation systems. However, most of these methods focus on single-objective optimization, fail to fully integrate multi-source heterogeneous data, and lack consideration of the dynamic coordination of users' short-term interests and long-term preferences, resulting in limited real-time and diversity of recommendation results.

In recent years, recommendation methods based on deep learning and knowledge graphs have gradually become a research hotspot. For example, Li et al. [9] used a deep learning method with nonlinear learning capabilities to automatically mine the potential relationship between users and products. Kim et al. [10] proposed a model that integrates local and global features to efficiently analyze review texts and extract user preferences. Liu et al. [11] proposed a structured framework and combined it with deep learning for embedded learning, prediction and reasoning, and verified its effectiveness through experiments in smart home product design prediction. However, existing methods still have obvious shortcomings in terms of computational efficiency, real-time adaptability and comprehensive optimization of user experience in complex scenarios. This paper proposes a personalized recommendation framework that integrates dynamic knowledge graphs, causal inference and deep reinforcement learning to achieve more accurate and personalized user experience improvement.

3. Method

3.1 Dynamic e-Commerce Knowledge Graph Construction

The shift in user interests, the rapid release of new products, and the shift in seasonal hot spots all require the underlying data representation to respond to these changes in real time. Therefore, this study constructed a dynamic e-commerce knowledge graph, the structure of which is shown in Figure 1.

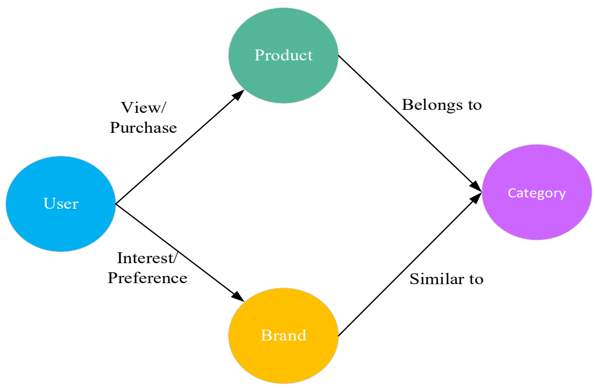

Figure 1: Schematic diagram of building a dynamic e-commerce knowledge graph

Note: The entity nodes in the graph include users, products, brands, and categories, and the edges represent interactions and semantic relationships.

First, multi-source data fusion and entity relationship definition are performed. This paper integrates multiple types of data from business databases, including product metadata (such as brand, category, attributes), user portraits (such as basic demographics), and the most critical real-time user behavior sequences (click, add to cart, purchase, etc.). On this basis, this paper clearly defines the entity types (such as user, product, brand, category) and relationship types (such as "purchase", "browse", "belong to", "similar to") in the graph. Unlike static graphs, this solution particularly emphasizes the dynamic properties of relationships [12].

Next, a dynamic relationship weight mechanism based on time decay is designed. For example, the browsing relationship weight $w_{ui}^{view}(t)$ between user u and item i at a certain time t is recalculated using a time decay function (such as an exponential decay function) based on the interval between the user u's most recent browsing time and the current time and the total number of historical browsing times, as shown in Formula (1).

$$w_{ui}^{view}(t)=\frac{1}{1+\lambda\cdot(t-t_{last})}\cdot count(u,i) \tag{1}$$

t_{last} is the time of user u's most recent interaction with item i, λ is the decay factor, and count(u,i) represents the number of historical interactions between user u and item i. This enables the graph to automatically deemphasize users' outdated interests and reinforce emerging preferences reflected by their recent behavior, thereby more accurately depicting the dynamic changes in user intent.

Finally, it implements real-time incremental updates and graph storage. To ensure real-time recommendations, it designs an incremental update pipeline. By monitoring the user's real-time behavior message stream, Kafka, the system triggers real-time weight recalculation and structural updates of the corresponding subgraphs, rather than performing a full reconstruction. Assuming that the current graph update process is G_t and the graph before the update is G_{t-1}, the incremental update process can be expressed as (2):

$$G_t=G_{t-1}\cup\Delta G_t \tag{2}$$

ΔG_t represents the graph changes introduced by this incremental update, and G_t represents the updated graph. The updated graph information is stored in the Neo4j graph database for efficient traversal and semantic query by downstream recommendation modules.

3.2 User Interest Debiasing Processing Based on Causal Inference

The historical interaction data that recommendation systems rely on is not an unbiased reflection of users' true preferences, but rather is heavily confounded by factors such as exposure mechanisms and platform interface design. To address this challenge at the source, this study introduces a causal inference framework. The core of this approach is to model user click behavior on products as a causal problem. The core idea is that user click behavior is driven by two factors: the user's true preference for the product (the causal effect) and the product's exposure and inherent popularity (the confounding bias).

First, this paper constructs a counterfactual reasoning model to estimate the exposure probability of a product. For any user-product pair (u, i), this paper uses its contextual features (such as the user's historical activity, the product's inherent attributes and real-time popularity) to train an exposure model to predict the probability $P(O=1|u,i,C)$ of the product being exposed to user u at the current time, that is, the probability that product i is seen by user u. The probability is calculated as follows (3):

$$P(O=1|u,i,C)=\sigma\left(\theta^T C\right) \tag{3}$$

σ is the Sigmoid activation function, θ is the model parameter, and C is the context feature vector of the user and the product. This probability value quantifies the contribution of other factors (such as operating position and popularity) to this interaction in addition to the user's real interest [13].

Next, this paper uses the Inverse Propensity Weighting (IPW) technique to correct the training samples. During the model training phase, for an observed click record, this paper no longer regards it as an ordinary positive sample, but assigns it a weight, which is the inverse of its exposure probability $P(O=1|u,i,C)$, expressed as (4):

ISSN 2616-5902 Vol. 7, Issue 11: 134-140, DOI: 10.25236/AJBM.2025.071118

$$w_{ui}=\frac{1}{P(O=1|u,i,C)} \tag{4}$$

In this way, long-tail products that are clicked despite having few exposure opportunities (low exposure probability) are assigned a higher weight w_{ui} during training. Conversely, popular products that are clicked only because they are frequently seen (high exposure probability) are assigned a correspondingly lower weight. This shifts the focus of model training from "fitting biased observational data" to "estimating unbiased user interests."

3.3 Deep Reinforcement Learning Strategy Optimization with ICM

This study designed a deep reinforcement learning strategy that incorporates an Intrinsic Curiosity Module (ICM) to achieve the ultimate goal of optimizing the long-term user experience. This framework models the recommendation process as a sequential decision-making task: at each time step, the agent (the recommender system) selects an action (i.e., generates a list of recommendations) based on the current environment state. The environment (the user) then generates feedback and updates its state. The agent's goal is to learn a strategy that maximizes the long-term cumulative reward.

First, a state representation is constructed that integrates information from multiple sources. The design of the state space S is fundamental to policy learning. The state vector in this study is not simply a user ID embedding, but rather incorporates three pieces of information: 1) a user's real-time, short-term interest session representation extracted from a dynamic knowledge graph; 2) a user's long-term, stable interest profile after causal debiasing; and 3) current contextual information (such as time and access channel). This rich state representation provides the agent with comprehensive and accurate decision-making support.

Secondly, it designs a composite reward function that incorporates intrinsic curiosity. This is the core innovation of this solution. The reward function $R(s,a)$ is no longer just an external commercial reward such as clicks or purchases. This paper introduces ICM, which quantifies the novelty of the agent's actions through a self-supervised prediction model. Formula (5)

$$R_{intrinsic}(s,a)= // \hat{s}_{next}(s,a)-s_{next}(s,a) // \tag{5}$$

$\hat{s}_{next}(s,a)$ represents the next state predicted by the agent after executing action a in state s, while $s_{next}(s,a)$ is the actual next state obtained. Specifically, ICM predicts the state change after executing action a in state s and uses the prediction error as the intrinsic reward (Intrinsic Reward) - the more difficult it is for the agent to predict the consequences of its behavior (i.e., exploring unknown areas), the higher the exploration reward obtained. Finally, the proximal policy optimization (PPO) algorithm is used for stable training [14]. The PPO algorithm can effectively constrain the step size of the policy update through its clipped surrogate objective function, avoiding drastic fluctuations and performance crashes during training. Its policy function is as follows (6).

$$L^{CLIP}(\theta)=\hat{E}_t\left[\min\left(r_t(\theta)\hat{A}_t, \mathrm{clip}\left(r_t(\theta),1-\epsilon,1+\epsilon\right)\hat{A}_t\right)\right] \tag{6}$$

$r_t(\theta)$ is the policy ratio, $\hat{A}_t$ is the advantage estimate, ϵ is the clipping parameter, and $\hat{E}_t$ is the expected value. To ensure the reproducibility of the experiment and the robustness of the policy optimization process, this study fine-tuned the key hyperparameters of the deep reinforcement learning model, as shown in Table 1.

Table 1: Key hyperparameter configurations of the strategy model

Category	Parameter Name	Value
Reinforcement Learning	Discount Factor (gamma)	0.99
	Trajectory Length (T)	20
Network Architecture	Actor / Critic Network	MLP (256,128)
Optimizer & Learning Rate	Adam	0.0003
Exploration Strategy	ICM Intrinsic Reward Weight (beta)	0.1
Training Settings	PPO Clipping Range (epsilon)	0.2
	Batch Size	512
	Epochs	100

ISSN 2616-5902 Vol. 7, Issue 11: 134-140, DOI: 10.25236/AJBM.2025.071118

4. Results and Discussion

4.1 Experimental Setup

To comprehensively and rigorously validate the effectiveness of the innovative framework proposed in this study, this evaluation will utilize a comprehensive validation approach combining online A/B testing, offline evaluation, and simulated environment experiments. All experiments are based on a desensitized dataset extracted from a large domestic e-commerce platform. This dataset contains anonymized behavior logs of approximately 2 million users over the past six months, including clicks, purchases, and add-to-cart information, as well as product attribute information. The comparison models (baselines) are:

NCF: Neural Collaborative Filtering, a representative collaborative filtering model.

DIN: Deep Interest Network, used to evaluate a model's ability to capture dynamic user interests.

DDPG: A reinforcement learning model that uses a deep deterministic policy gradient algorithm and whose reward function only includes business metrics. This is used for comparison and verification of the advantages of the composite reward function in this paper.

4.2 Experimental Analysis

(1) Long-term User Value Enhancement Verification Experiment

This experiment aims to evaluate the effectiveness of the new recommendation algorithm in improving long-term user value. This paper designs a four-week online A/B test, randomly dividing users into an experimental group and a control group, using the new algorithm and the traditional algorithm respectively. Figure 2 shows the behavioral data of the two groups of users monitored, focusing on the two core indicators of average session duration and seven-day retention rate to quantify the actual impact of algorithm optimization on user engagement and loyalty.

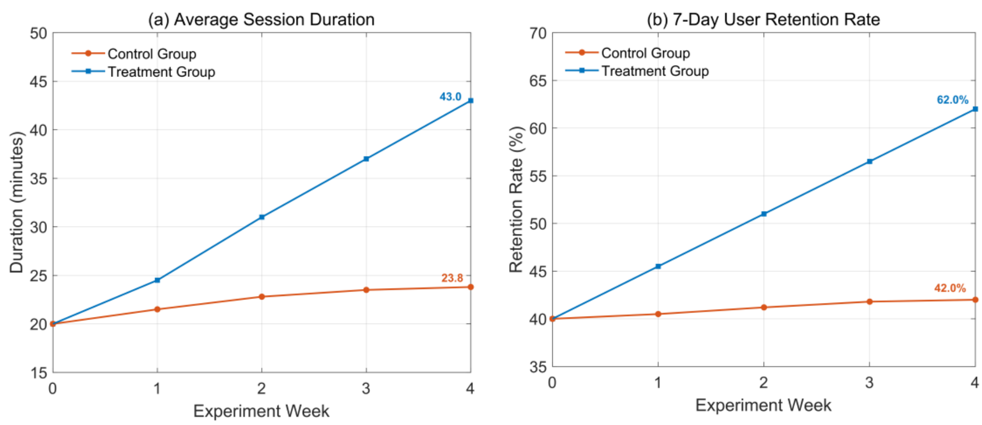

Figure 2: Long-term user value enhancement verification evaluation

Figures 2 (a–b) show the trends in average session duration and seven-day user retention for the experimental and control groups, respectively, over different weeks. After a four-week A/B test, the average session duration for the experimental group reaches 43 minutes, an increase of approximately 80.7% compared to the control group (23.8 minutes). Furthermore, the seven-day user retention rate for the experimental group stabilizes at 62%, a significant increase of 20 percentage points compared to the control group (42%). This data demonstrates that the algorithm, by accurately capturing users' dynamic interests and exploring potential needs, effectively enhances user stickiness and platform loyalty, achieving a shift from short-term efficiency to long-term user value optimization.

(2) Effectiveness Experiment on Recommendation System Bias Correction

This experiment uses an offline testing method to evaluate the effectiveness of recommendation algorithms in correcting exposure bias. By constructing a historical dataset containing known biases, the performance of a traditional baseline model and an innovative model based on causal inference in recommending long-tail products is compared. Figure 3 shows the changes in exposure ratios of long-tail products for the two models, quantifying the actual effectiveness of the bias correction algorithm in improving recommendation diversity.

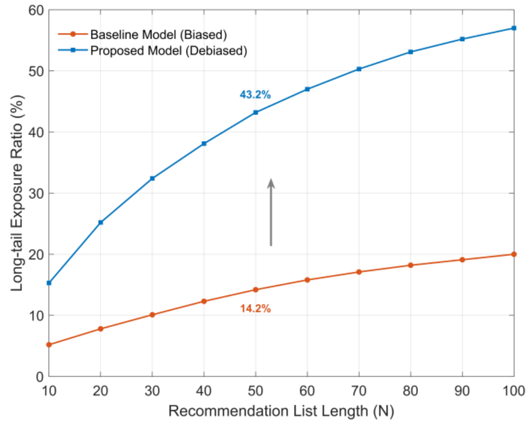

Figure 3: Evaluation of the effectiveness of bias correction in recommendation systems

As shown in Figure 3, when the recommendation list length N = 50, the new algorithm achieves a 43.2% exposure rate for long-tail products, significantly outperforming the traditional model's 14.2%, an improvement of 29.0 percentage points. This demonstrates that the algorithm effectively removes exposure bias in historical data through causal inference technology, enabling more accurate identification and recommendation of potentially high-quality long-tail products. This significantly improves the diversity and fairness of the recommendation system, providing an effective solution for mining the value of long-tail inventory on e-commerce platforms.

(3) Exploration - Optimizing Experiments Using Balanced Strategies

This experiment, conducted in a simulated environment, aims to verify the effectiveness of the designed composite reward function, incorporating the Intrinsic Curiosity Module (ICM), in balancing the exploration-exploitation problem. Table 2 compares a reinforcement learning strategy using a traditional single commercial reward with the composite reward strategy proposed in this study. The performance of the strategy in uncovering users' potential interests is quantified by evaluating the success rate of exploring new categories over multiple training cycles.

Table 2: Optimization evaluation of exploration-exploitation balance strategy

Training Epoch	Baseline Exploration Success Rate(%)	Proposed Exploration Success Rate(%)
1	12.5	15.8
2	16.3	28.4
3	18.1	39.7
4	19.5	47.2

As shown in Table 2, after four training cycles, the compound reward strategy achieves a 47.2% success rate in exploring new categories, a 142% improvement compared to the 19.5% success rate achieved with the traditional strategy. This result demonstrates that the designed reward function, incorporating an intrinsic curiosity-driven exploration mechanism, can more effectively guide the agent out of local optima and proactively discover unmet user interests, successfully resolving the core trade-off between exploration and exploitation in recommendation systems.

4.3 Experimental Discussion

Combining the results of the three experiments, it can be seen that the personalized recommendation framework proposed in this study has shown significant advantages in terms of user value enhancement, system bias correction, and exploration and utilization balance. The long-term user value enhancement experiment shows that the algorithm can effectively extend the user session duration and improve the retention rate, thereby enhancing the platform's stickiness and loyalty; the bias correction experiment further proves that by removing the exposure bias in historical data through causal inference, the model can significantly increase the exposure and clicks of long-tail products, improving the diversity and fairness of recommendation results; and the exploration and utilization balance experiment shows that the compound reward function that integrates the intrinsic curiosity mechanism can significantly improve the exploration success rate of new categories, effectively guide the intelligent agent to jump out of the local optimum, and explore the potential interests of users. The three aspects of evidence together show that this framework not only solves the limitations of traditional recommendation systems under the short-term accuracy orientation, but also provides a practical path for the platform to achieve long-term value growth and healthy ecological development.

5. Conclusion

This paper addresses the homogeneity and lack of long-term value in e-commerce recommendation systems and proposes a personalized recommendation framework that integrates dynamic knowledge graphs, causal inference, and reinforcement learning. The results show that this approach not only effectively improves user stickiness and retention, but also achieves good results in correcting system biases, enhancing the exposure of long-tail products, and improving the balance between exploration and utilization, verifying its feasibility in optimizing user experience and platform ecology. However, this study still has shortcomings. For example, there is still a gap between the experimental environment and real business scenarios, and the performance and stability of the model in large-scale real-time applications need further testing. In the future, it is possible to combine multimodal data to expand user portraits, explore the portability of cross-domain recommendations, and verify them in a larger-scale platform environment to promote the further development of recommendation systems towards diversification, fairness, and long-term value orientation.

Acknowledgements

Research Project of Changchun Humanities and Sciences College: "A Study on the Synergetic Mechanism of Enterprise Management Innovation and Industrial Upgrading under the Background of Digital Transformation". Project No. 2025KG05

References

[1] Zhang Y. The application of e-commerce recommendation system in smart cities based on big data and cloud computing[J]. Computer Science and Information Systems, 2021, 18(4): 1359-1378.
[2] Gulzar Y, Alwan A A, Abdullah R M, et al. OCA: ordered clustering-based algorithm for e-commerce recommendation system[J]. Sustainability, 2023, 15(4): 2947-2951.
[3] Loukili M, Messaoudi F, El Ghazi M. Machine learning based recommender system for e-commerce[J]. IAES International Journal of Artificial Intelligence, 2023, 12(4): 1803-1811.
[4] Shankar A, Perumal P, Subramanian M, et al. An intelligent recommendation system in e-commerce using ensemble learning[J]. Multimedia Tools and Applications, 2024, 83(16): 48521-48537.
[5] Wu K, Chi K. Enhanced e-commerce customer engagement: A comprehensive three-tiered recommendation system[J]. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 2023, 2(3): 348-359.
[6] Wei Fen. Research on design and performance optimization of online recommendation system based on machine learning algorithm[J]. Information and Computer, 2024, 36(3):89-91.
[7] Necula S C, Păvăloaia V D. AI-driven recommendations: A systematic review of the state of the art in e-commerce[J]. Applied Sciences, 2023, 13(9): 5531-5542.
[8] Kim J, Choi I, Li Q. Customer satisfaction of recommender system: Examining accuracy and diversity in several types of recommendation approaches[J]. Sustainability, 2021, 13(11): 6165-6172.
[9] Li W, Cai Y, Hanafiah M H, et al. An empirical study on personalized product recommendation based on cross-border e-commerce customer data analysis[J]. Journal of Organizational and End User Computing (JOEUC), 2024, 36(1): 1-16.
[10] Kim N, Lim H, Li Q, et al. Enhancing Review-Based Recommendations Through Local and Global Feature Fusion[J]. Electronics, 2025, 14(13): 2540-2551.
[11] Liu A, Zhang D, Wang Y, et al. Knowledge graph with machine learning for product design[J]. CIRP Annals, 2022, 71(1): 117-120.
[12] Behera G, Nain N. Collaborative filtering with temporal features for movie recommendation system[J]. Procedia Computer Science, 2023, 218(1): 1366-1373.
[13] Bączkiewicz A, Kizielewicz B, Shekhovtsov A, et al. Methodical aspects of MCDM based E-commerce recommender system[J]. Journal of Theoretical and Applied Electronic Commerce Research, 2021, 16(6): 2192-2229.
[14] Bawack R E, Wamba S F, Carillo K D A, et al. Artificial intelligence in E-Commerce: a bibliometric study and literature review[J]. Electronic markets, 2022, 32(1): 297-338.