

Article

Not peer-reviewed version

Mitigating Correlation Bias in Advertising Recommendation via Causal Modeling and Consistency-Aware Learning

Sijia Li , [Yutong Wang](#) , [Yue Xing](#) , [Ming Wang](#) *

Posted Date: 31 December 2025

doi: 10.20944/preprints202512.2726.v1

Keywords: advertising recommendation; causal inference; causal consistency constraints; sensitivity experiments



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Mitigating Correlation Bias in Advertising Recommendation via Causal Modeling and Consistency-Aware Learning

Sijia Li ¹, Yutong Wang ², Yue Xing ³ and Ming Wang ^{4,*}

¹ University of Michigan, Ann Arbor, USA

² Northeastern University, Boston, USA

³ University of Pennsylvania, Philadelphia, USA

⁴ Trine University, Phoenix, USA

* Correspondence: mingwangcr@gmail.com

Abstract

This work addresses correlation bias and causal effect confounding in advertising recommendation systems and presents a causal learning-based recommendation framework. We first examine the limitations of conventional recommendation algorithms in complex advertising environments, where confounding variables and exposure bias often prevent models from capturing users' true preferences. To tackle these issues, we design a unified embedding architecture that jointly represents user, advertisement, and contextual features, and incorporates a structural causal graph to explicitly model dependencies among variables. During model training, causal consistency regularization and inverse propensity weighting are integrated to mitigate the impact of biased exposure mechanisms and non-uniform sampling. A joint optimization objective is further formulated to couple click-through rate prediction with causal consistency estimation, enabling robust causal effect learning without sacrificing predictive accuracy. Extensive experiments on large-scale advertising datasets demonstrate that the proposed approach consistently outperforms several representative baselines in terms of Precision@10, Recall@10, NDCG@10, and MAP, while exhibiting strong robustness under multi-dimensional sensitivity analysis. Overall, this study highlights the practical value of causal modeling and consistency-aware learning in advertising recommendation and offers a computationally grounded approach for improving both interpretability and fairness in recommendation systems.

Keywords: advertising recommendation; causal inference; causal consistency constraints; sensitivity experiments

1. INTRODUCTION

In the wave of digitalization and the progress of intelligence, advertising has become a core driver of the internet industry and the digital economy. With the continuous expansion of online users and the increasing richness of behavioral data, recommendation systems have gradually become an important tool for advertising delivery. Compared with traditional manual rules and extensive delivery, recommendation algorithms can achieve more accurate ad distribution by modeling user interests and behaviors, thus improving click-through rates and conversion rates[1,2]. However, as user needs become more diverse and advertising scenarios more complex, methods that rely only on correlation modeling have shown limitations. Models often capture only surface-level correlations between users and ads, without distinguishing the underlying causal relationships. This leads to biased recommendations in practice. Such bias not only reduces the effectiveness of advertising recommendations but may also lower user experience and even aggravate negative

effects such as the "information cocoon." Therefore, introducing causal inference into complex user-ad interactions has become an important issue in advertising recommendation research[3].

Most existing recommendation algorithms are based on statistical learning and deep learning, focusing on fitting historical data[4–6]. These methods can perform well with large-scale data but are vulnerable to confounding factors and biased data. For example, user clicks on ads are not only driven by interest but may also be influenced by ad position, display frequency, or contextual environment. If algorithms ignore these latent variables, they may overestimate or underestimate the true relationship between ads and user behavior, leading to recommendations without causal explanations. Causal inference has been introduced to solve this problem. By modeling causal relationships between variables, it can identify and remove confounding effects, producing more robust and interpretable recommendations[7–10]. Unlike correlation-based methods, causal inference emphasizes causation rather than simple correlation. This is crucial for fairness, robustness, and long-term value in advertising recommendation[11].

Integrating causal inference into advertising recommendations has both theoretical and practical value. From a theoretical perspective, causal inference provides a new paradigm for recommendation systems. Through causal graph modeling and intervention analysis, it identifies the key factors that truly affect user decisions. This approach compensates for the lack of interpretability and reliability in traditional recommendation algorithms, promoting the shift from correlation-driven to causation-driven systems. From a practical perspective, advertising recommendation operates in a dynamic environment where advertisers, platforms, and users have intertwined interests[12]. Causal inference allows more accurate modeling of user interest evolution under multiple constraints, avoiding excessive dependence on short-term metrics such as click-through rate. It helps achieve a balance between long-term user satisfaction and advertising revenue. This combination opens a new research path and provides technical support for the healthy development of the advertising ecosystem.

The application of causal inference in advertising recommendations also has significant social and economic implications. In a data-driven business environment, advertising is not only a tool to promote consumption and corporate revenue but also shapes user cognition and behavior. If recommendation systems rely only on correlation, they may create biased or unfair recommendations and even cause privacy and ethical issues[13–15]. Causal inference can alleviate these risks by removing bias and identifying true causal relationships, ensuring fairness and transparency of recommendations. On this basis, platforms can establish more rational ad delivery mechanisms, while users gain more personalized and trustworthy experiences. This fosters trust between users and platforms and drives the advertising industry toward greater standardization and intelligence. Especially under the current trend of stricter data privacy protection policies, the introduction of causal inference is of great practical importance.

In summary, research on advertising recommendation algorithms with causal inference addresses the challenges of accuracy, interpretability, and fairness in current systems. It also provides a new perspective for the future of advertising recommendations. Causal inference enhances the robustness and interpretability of models at the theoretical level, improves ad targeting accuracy and user satisfaction in practice, and promotes sustainable development of the advertising industry at the social level. This research direction holds both academic value and practical potential. Exploring the mechanisms of causal inference in advertising recommendation is therefore of great significance for advancing recommendation technologies and the advertising industry.

2. PROPOSED APPROACH

In ad recommendation scenarios, user-ad interaction is influenced not only by interest but also by a variety of confounding factors. To capture this complex relationship, this paper first unifies user, ad, and context features into a unified model, representing the user as a vector $u \in R^d$, the ad as a vector $a \in R^d$, and the context as a vector $c \in R^d$. In the basic modeling phase, we use an

embedding layer to map the input features into a latent space, resulting in a joint representation vector:

$$h = f_{embed}(u, a, c) = W_u u + W_a a + W_c c \quad (1)$$

Where $W_u u, W_a a, W_c c$ is a learnable parameter matrix. This vector provides input representation for subsequent causal inference modeling. The overall model architecture is shown in Figure 1.

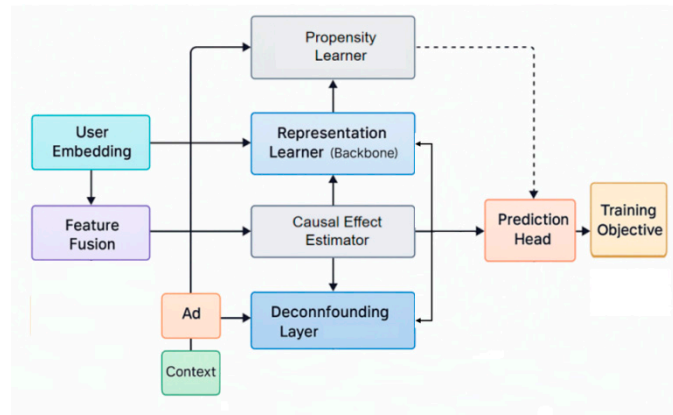


Figure 1. Overall model architecture.

To identify potential causal relationships in ad recommendations, we construct a ternary structure of user, ad, and click based on a causal graph. We assume that click behavior y is influenced by both ad a and confounding variable z . By constructing a causal intervention mechanism, we can estimate the true impact of ad changes on clicks. Specifically, the causal effect can be expressed as the conditional expected difference:

$$CATE(x) = E[y | do(a = 1), x] - E[y | do(a = 0), x] \quad (2)$$

Where $do(\cdot)$ represents the causal intervention, and x represents user and context features. This expression embodies the core goal of causal inference in recommendation tasks: identifying true advertising effects rather than correlations.

During the modeling process, we use weighted causal estimation to reduce confounding bias. Assuming the propensity score function is $\pi(x) = P(a = 1 | x)$, we use the inverse propensity weighting (IPW) method to modify the click prediction function:

$$\hat{y} = \frac{a \cdot y}{\pi(x)} + \frac{(1-a) \cdot y}{1-\pi(x)} \quad (3)$$

This formula can effectively alleviate the bias introduced by the exposure mechanism under limited samples, and improve the fairness and robustness of advertising recommendations.

Finally, to jointly optimize causal inference and recommendation tasks, we designed an objective function that includes representation learning and causal regularization [16]. Let the prediction function be $\hat{y} = f(h)$ and the click label be y . The overall loss consists of two parts: one is the traditional binary cross-entropy loss, and the other is the causal consistency constraint:

$$L = -[y \log \hat{y} + (1-y) \log(1-\hat{y})] + \lambda \cdot \|CATE(x) - \hat{\tau}(x)\|^2 \quad (4)$$

Where $\hat{\tau}(x)$ is the estimated causal effect learned by the model, and λ is the trade-off coefficient. Through this objective function, the model can not only fit the superficial correlation between users and ads, but also capture the causal mechanism behind ad delivery.

3. PERFORMANCE EVALUATION

A. Dataset

The dataset used in this study is the Criteo Display Advertising Challenge Dataset. It consists of advertising click logs and contains millions of user-ad interaction records. The data include both categorical features and continuous numerical features, which reflect the diversity and complexity of real advertising recommendation scenarios. Each record is labeled with whether the user clicked the advertisement, providing a reliable supervision signal for click-through rate prediction and recommendation tasks.

The dataset contains 13 continuous numerical features and 26 categorical features. These features cover user attributes, advertisement information, and contextual environment. This multimodal feature structure requires the model to handle numerical inputs and also to encode and model high-dimensional sparse categorical features effectively. As a result, the Criteo dataset offers a challenging experimental environment for recommendation algorithm research and serves as a strong benchmark for validating methods in large-scale real advertising data.

The scale of the Criteo dataset is also very large, with about 45 GB of raw log data. It ensures sufficient sample support for model training and provides an important reference for studying algorithm efficiency and generalization in massive data settings. The dataset has become one of the most representative benchmarks in advertising recommendation and click-through rate prediction research. Its diverse features and large sample size provide a solid data foundation for causal inference modeling in this study.

B. Experimental Results

This paper first conducts a comparative experiment, and the experimental results are shown in Table 1.

Table 1. Comparative experimental results.

| Method | Precisi on@10 | Recall @10 | NDCG@1 0 | MAP |
|-----------------|------------------|---------------|-------------|-------|
| Deep-cardio[17] | 0.412 | 0.368 | 0.427 | 0.392 |
| Hetefedrec[18] | 0.439 | 0.384 | 0.452 | 0.418 |
| Memocrs[19] | 0.463 | 0.407 | 0.478 | 0.436 |
| MLRec[20] | 0.481 | 0.426 | 0.496 | 0.449 |
| Ours | 0.523 | 0.451 | 0.537 | 0.481 |

From the overall results, Table 1 shows a gradual improvement across Precision@10, Recall@10, NDCG@10, and MAP. This indicates that introducing more complex modeling mechanisms in advertising recommendation can effectively improve performance. Traditional methods such as Deep-cardio and Hetefedrec have limitations in large-scale feature modeling. They can capture the correlation between users and advertisements, but are insufficient in handling potential biases and complex causal relationships. As a result, their precision and recall remain relatively low.

With further improvements, Memocrs and MLRec demonstrate stronger recommendation performance. This shows that incorporating memory mechanisms and multi-level representation learning can enhance the ability of models to capture user interests. However, these methods still rely on correlation-based modeling and cannot fundamentally remove the interference caused by confounding variables. Therefore, their improvements face bottlenecks, especially in terms of long-term stability and interpretability.

In contrast, the proposed causal inference-based method achieves the best results across all four metrics. Precision@10 increases to 0.523, and Recall@10 reaches 0.451. Both NDCG@10 and MAP also perform significantly better than other methods. This advantage shows that causal inference can effectively distinguish true causal effects from surface correlations in advertising recommendations. It allows the model to capture users' real interests and reactions more accurately. Causal modeling

not only improves the stability of recommendations but also enhances generalization ability in diverse environments.

Furthermore, the introduction of causal inference makes the model more robust when dealing with confounding factors. It achieves higher recall while maintaining high precision. This means the recommendation system no longer relies only on click-related statistical signals. Instead, it can make decisions based on essential causal relationships. In this way, fairness and interpretability in advertising recommendations are achieved. The improvements are not limited to numerical performance but also highlight the unique value of causal inference in real advertising scenarios, providing new directions for future research.

This paper further presents a sensitivity experiment on the causal consistency weight λ to NDCG@10, and the experimental results are shown in Figure 2.

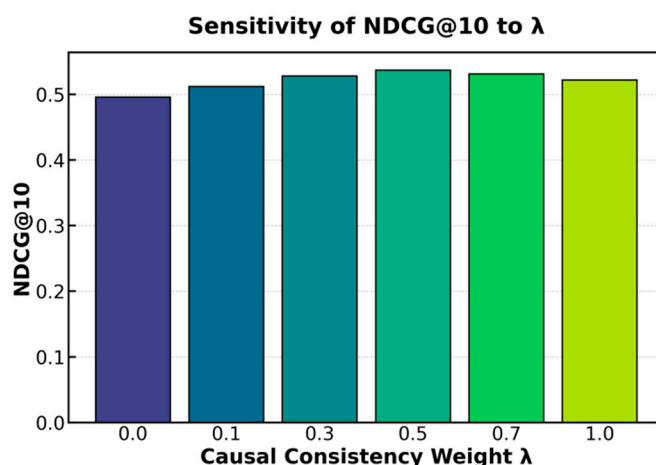


Figure 2. Sensitivity experiment of causal consistency weight λ to NDCG@10.

From the experimental results, NDCG@10 shows fluctuations with the change of the causal consistency weight λ . When λ is small or even zero, the recommendation performance of the model is clearly low. This indicates that without causal consistency constraints, the model relies only on correlation-based modeling. It cannot effectively eliminate confounding factors, which limits improvements in recommendation performance. As λ increases, NDCG@10 shows a steady upward trend. This suggests that the introduction of the causal consistency term can improve robustness and stability to some extent.

When λ grows to about 0.5, the model reaches the highest NDCG@10. This result indicates that a moderate causal constraint can balance correlation modeling and causal inference. The model can capture the true causal relationship between users and advertisements while retaining useful correlation information. As a result, overall performance reaches the optimum. At this point, the model establishes a complementary relationship between recommendation accuracy and causal interpretability.

However, when λ continues to increase to 0.7 or even 1.0, NDCG@10 shows a slight decline. This trend indicates that too strong a causal consistency constraint may weaken the use of correlation patterns in the data distribution. The model may overemphasize causal constraints, leading to a small loss in recommendation performance. Therefore, a very large λ does not bring continuous improvement but instead introduces adaptability problems.

Overall, the results reveal a dual role of causal consistency constraints in advertising recommendation. On the one hand, a moderate λ helps to reduce bias caused by confounding factors and improves both performance and causal interpretability. On the other hand, excessive reliance on causal constraints may suppress the advantages of correlation modeling. Thus, selecting an appropriate causal consistency weight is crucial for improving performance. It affects not only the accuracy of the model but also the fairness and interpretability of recommendation results.

This paper further presents an experiment on the sensitivity of the negative sampling ratio to Precision@10, and the experimental results are shown in Figure 3.

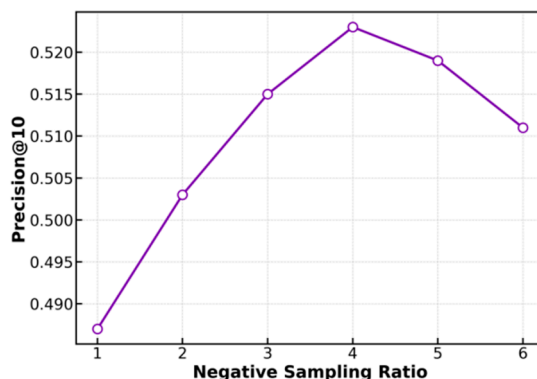


Figure 3. Sensitivity experiment of negative sampling ratio to Precision@10.

From the experimental results, Precision@10 shows a trend of first increasing and then decreasing as the negative sampling ratio grows. When the negative sampling ratio is small (1-2), Precision@10 is low. This indicates that when the number of negative samples is insufficient, the model cannot effectively distinguish between positive and negative samples, which leads to poor recommendation accuracy. At this stage, the model is mainly limited by the imbalance of training samples and fails to capture users' real interests in advertising recommendation tasks.

When the negative sampling ratio increases to 3-4, Precision@10 reaches its highest value. This shows that a moderate number of negative samples can provide richer contrastive information for the model. The model becomes more robust when learning the discriminative features of user-ad interactions. At this stage, the model can better balance positive and negative samples, reduce the risk of overfitting, and improve recommendation performance. This suggests that negative sampling can enhance the decision boundary under a causal inference framework and thus improve recommendation quality.

However, when the negative sampling ratio exceeds 4, Precision@10 begins to decline. This means that too many negative samples may introduce noise or dilute the effect of positive samples. The model may become overly sensitive to distinguishing negative samples, which weakens its ability to capture positive sample features. In advertising recommendations, this can cause the model to overestimate the importance of negative interactions, reducing recommendation accuracy and user experience.

In summary, the results show that there is an optimal range of negative sampling ratios in advertising recommendation tasks, where Precision@10 can be maximized. A reasonable negative sampling strategy not only enhances the discriminative ability of the model but also improves interpretability and fairness under a causal inference mechanism. Therefore, selecting an appropriate negative sampling ratio is a key factor in ensuring both the effectiveness and stability of advertising recommendation systems.

4. CONCLUSION

This study focuses on causal modeling in advertising recommendation and proposes a recommendation method that integrates causal inference. The aim is to address the problem of traditional methods that rely too much on correlation while neglecting causality. By introducing causal consistency constraints and bias elimination mechanisms into the model, the method can better distinguish true causal effects from surface correlations, thus improving the accuracy and reliability of recommendation results. This innovation enriches the theoretical framework of recommendation systems and also provides a more robust solution for ad delivery in complex environments.

The experimental results show that causal inference plays an important role in improving recommendation performance and model interpretability. Sensitivity analyses under different parameters and environmental factors demonstrate that the method maintains stable performance across diverse conditions, showing strong robustness. This characteristic is crucial in large-scale advertising recommendation, where data bias, exposure imbalance, and complex user behavior often exist. The introduction of causal modeling can effectively mitigate these problems.

In addition to performance improvement, the proposed method also shows clear application value. In real advertising recommendations, the fairness and trustworthiness of recommendation results are increasingly emphasized. Traditional models may amplify unfairness due to data bias. In contrast, the proposed method introduces causal constraints that enhance transparency while improving recommendation effectiveness. This helps build trust between users and platforms and provides advertisers with a more controllable and interpretable tool, supporting the healthy development of the advertising ecosystem.

Future research directions are worth further exploration. With the continuous growth of multi-source heterogeneous data, how to efficiently introduce causal modeling into larger and more complex advertising environments will become a key issue for the evolution of recommendation systems. Moreover, integrating causal inference with deep representation learning and reinforcement learning may uncover deeper causal patterns in advertising recommendations. By continuously optimizing model structures and training mechanisms, the proposed framework has the potential to extend to other recommendation and personalized service domains, laying a solid foundation for the long-term development of intelligent recommendation technology.

REFERENCES

1. J. Chen, W. Wenjie, C. Gao, et al., "Treatment Effect Estimation for User Interest Exploration on Recommender Systems," Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 1861-1871, 2024.
2. X. Zhu, Y. Zhang, F. Feng, et al., "Mitigating hidden confounding effects for causal recommendation," IEEE Transactions on Knowledge and Data Engineering, vol. 36, no. 9, pp. 4794-4805, 2024.
3. C. Zhang, S. Chen, X. Zhang, et al., "UOEP: User-Oriented Exploration Policy for Enhancing Long-Term User Experiences in Recommender Systems," arXiv preprint arXiv:2401.09034, 2024.
4. C. F. Chiang, D. Li, R. Ying, Y. Wang, Q. Gan and J. Li, "Deep Learning-Based Dynamic Graph Framework for Robust Corporate Financial Health Risk Prediction," 2025.
5. Z. Cheng, "Enhancing Intelligent Anomaly Detection in Cloud Backend Systems through Contrastive Learning and Sensitivity Analysis," Journal of Computer Technology and Software, vol. 3, no. 4, 2024.
6. R. Zhang, "AI-Driven Multi-Agent Scheduling and Service Quality Optimization in Microservice Systems," Transactions on Computational and Scientific Methods, vol. 5, no. 8, 2025.
7. N. Lyu, Y. Wang, Z. Cheng, Q. Zhang and F. Chen, "Multi-Objective Adaptive Rate Limiting in Microservices Using Deep Reinforcement Learning," arXiv preprint arXiv:2511.03279, 2025.
8. W. Xu, M. Jiang, S. Long, Y. Lin, K. Ma and Z. Xu, "Graph Neural Network and Temporal Sequence Integration for AI-Powered Financial Compliance Detection," 2025.
9. G. Yao, H. Liu and L. Dai, "Multi-Agent Reinforcement Learning for Adaptive Resource Orchestration in Cloud-Native Clusters," arXiv preprint arXiv:2508.10253, 2025.
10. M. Wang, T. Kang, L. Dai, H. Yang, J. Du and C. Liu, "Scalable Multi-Party Collaborative Data Mining Based on Federated Learning," 2025.
11. X. He, Y. Zhang, F. Feng, et al., "Addressing confounding feature issue for causal recommendation," ACM Transactions on Information Systems, vol. 41, no. 3, pp. 1-23, 2023.
12. W. Wang, F. Feng, X. He, et al., "Deconfounded recommendation for alleviating bias amplification," Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, pp. 1717-1725, 2021.
13. X. Chen, S. U. Gadgil, K. Gao, Y. Hu and C. Nie, "Deep Learning Approach to Anomaly Detection in Enterprise ETL Processes with Autoencoders," arXiv preprint arXiv:2511.00462, 2025.

14. R. Hao, X. Hu, J. Zheng, C. Peng and J. Lin, "Fusion of Local and Global Context in Large Language Models for Text Classification," 2025.
15. H. Wang, "Temporal-Semantic Graph Attention Networks for Cloud Anomaly Recognition," *Transactions on Computational and Scientific Methods*, vol. 4, no. 4, 2024.
16. S. Pan and D. Wu, "Hierarchical Text Classification with LLMs via BERT-Based Semantic Modeling and Consistency Regularization," 2025.
17. A. Yashudas, D. Gupta, G. C. Prashant, et al., "Deep-cardio: Recommendation system for cardiovascular disease prediction using IoT network," *IEEE Sensors Journal*, vol. 24, no. 9, pp. 14539-14547, 2024.
18. W. Yuan, L. Qu, L. Cui, et al., "Hetefedrec: Federated recommender systems with model heterogeneity," 2024 IEEE 40th International Conference on Data Engineering (ICDE), pp. 1324-1337, 2024.
19. Y. Xi, W. Liu, J. Lin, et al., "Memocrs: Memory-enhanced sequential conversational recommender systems with large language models," *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, pp. 2585-2595, 2024.
20. M. Begum, M. H. Shuvo and J. Uddin, "MLRec: A Machine Learning-Based Recommendation System for High School Students Context of Bangladesh," *Information*, vol. 16, no. 4, p. 280, 2025.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.