Dean&Francis

ISSN 2959-6157

Causal Inference for Recommender Systems: Methods, Challenges, and Applications

Yunlong Jiao[1, *]

[1]Department of Applied Math, Northeastern University, Boston, 02115, MA, the United States
*Corresponding author: jiao.yunl@northeastern.edu

Abstract:

Traditional recommendation systems based on user interaction data for correlation analysis have significant bias issues—such as selection bias, exposure bias, and popularity bias —that arise from the inherent limitations of observational data, making it difficult to predict user preferences accurately without bias and often leading to homogenized recommendations that fail to uncover users' potential interests. Users lose trust in recommendation systems because of this situation which makes it difficult to implement these systems in complex environments that need exact personalized healthcare services and intelligent decision systems. The review examines causal recommendation surveys and presents new approaches (debiasing and causal collaborative filtering and distillation) and demonstrates their applications in time-series analysis and healthcare and uplift modeling. The research demonstrates that causal inference methods remove confounding variables which produces improved prediction accuracy and better system operation understanding for users. Research on causal inference will progress through the combination of deep learning techniques which analyze intricate temporal relationships and multiple domains to create functional recommendation platforms. The research establishes vital foundations for upcoming investigations which will link causal inference to recommendation systems through its complete analysis of this field.

Keywords: Causal Inference; Recommender Systems; Counterfactual Learning; Deconfounding; Debiasing.

1. Introduction

The Internet now includes Recommender systems (RS) which deliver customized services to users throughout their entire life experience [1]. The three main categories of traditional systems consist of collaborative filtering and content-based filtering and hybrid systems which unite these two approaches [1]. The Internet-based analysis of user data including profiles and behavioral records and contextual elements makes recommender systems have become essential for business and entertainment and consulting and life services [2]. The algorithms used in practical applications succeed through user behavior data analysis yet they fail to discover the actual causal relationships which drive user behavior [3].

The correlation-based algorithm model produces three major problems which negatively impact model fairness and robustness and interpretability. The recommendation system faces three major biases which include selection bias and popularity bias and exposure bias that produce incorrect user preference models and reduce system generalization [4]. The system generates too many recommendations for popular items which causes it to ignore what users actually need. The combination of these issues with sparse data reduces the recommendation value for users who are new to the system and for products that have not been well represented [5]. The exposure bias occurs because users can only assess products they have encountered which creates non-random missing data. The current network environment enables traffic head effects to strengthen this bias which results in decreased diversity [6]. The inability of traditional algorithms to identify genuine causal relationships from fake correlations that result from confounding variables [2]. The predictive accuracy of the system decreases while producing social effects that include information cocoons and polarization [7].

The framework of causal inference solves all the issues which traditional algorithms encounter. The system uses causal relationship modeling to perform intervention-based reasoning which enables it to forecast unbiased user behavior responses to recommended content [8]. The shift from associative learning to causal modeling has led to major improvements in recommender system fairness and robustness and interpretability while creating research alignment with modern causal theory [9].

In the latest machine learning research, causal inference sets have become one of the most active frontier areas in recommender systems [10]. Causal inference provides a theoretical and algorithmic basis for solving the problems of bias, confusing variables and lack of interpretability that are common in traditional recommender algorithms [9]. Through frameworks such as potential outcome models and structural causal models (SCM), causal recommendation can not only solve the "what" question about user behavior, but also answer the "why" and "what if" questions about intervention measures [8].

A large number of academic studies and methodological innovations on causal inference have emerged. Gao et al. conducted a large-scale research survey, classified the main biases (including popularity bias, exposure bias and location bias) and their corresponding causal mitigation strategies, and finally described the future research trends [4]. Luo's team proposed a causal inference classification system, dividing causal inference into three frameworks: potential outcome paradigm, structural causal paradigm and counterfactual learning paradigm, providing a systematic overall perspective for causal recommendation research [5]. Zhu's research focuses on causal inference strategies that can mitigate bias, improve interpretability and generalization ability. It emphasizes that causal models can effectively improve the interpretability of recommendation systems by revealing stable and invariant causal relationships under different environments [11].

In addition to the above reviews on causal inference, many studies have further proposed targeted algorithmic frameworks. Wang et al. proposed a deconfounding recommender, which effectively eliminates hidden confounding factors by modeling potential exposure variables [2]. Zhang et al. extended this idea by modeling identifiable potential confounding factors, and achieved accurate recovery of real confounding factors with the help of identifiable variational autoencoders [9]. Xu et al. proposed causal collaborative filtering, which directly embeds causal relationships into the user-item interaction modeling process, thereby achieving more accurate and fair recommendations [10]. Zhang et al. further proposed causal distillation technology, which effectively alleviates the performance heterogeneity problem of recommendation systems by transferring causal knowledge between different models [8]. In addition, Liu et al. proposed a propensity score estimation method without random experiments, which successfully solved the problem of missing exposure information in recommendation scenarios [6].

The research paper conducts an extensive evaluation of recommender system causal inference through method explanations and challenge identification and multiple application demonstrations. The research investigates three main approaches to solve existing problems which affect current causal inference methods. The research performs an extensive review of current causal recommender system developments by using a single conceptual framework. The research explains how causal inference methods enhance recommender system performance through better bias reduction and improved interpretability and en-

ISSN 2959-6157

hanced generalization capabilities. The study determines critical research domains which will create unbiased and dependable and transparent recommendation systems through deep learning integration with causal inference. The research unites causal inference knowledge to establish a new research framework for recommenders while offering structured recommendations for its upcoming growth.

2. Development and Mainstream Techniques of Causal Recommender Systems

2.1 The Deconfounded Recommender

Traditional recommender systems use association learning to detect statistical patterns between users and items but this method fails to distinguish between causal relationships and correlational ones. The correlation-based modeling approach becomes susceptible to confounding bias because unmeasured variables which affect user-item interactions also influence user behavior in the future. The Deconfounded Recommender functions as the first causal framework which removes all factors that affect recommendation outcomes [2].

The system creates a „surrogate confounding factor“ through its multi-cause exposure model inference process. The model presents users with various items which function as causes while Z operates as a hidden variable to represent unmeasured confounding elements that affect their exposure patterns. The system learns Z through probabilistic decomposition methods which include Poisson decomposition and variational inference before applying it as a new covariate to study exposure effects on user feedback.

The resulting model can be represented as:

$$E\left[Y_{ui}\,\text{¨O}T_{ui},Z_u,X_{ui}\right]=f_\theta\left(T_{ui},Z_u,X_{ui}\right) \tag{1}$$

The model uses T_ui to represent exposure states and X_ui to represent observable features and Z_u to represent the inferred alternative confounding factor for user u. The model uses deconfounded data to estimate f_θ which enables it to predict the actual intervention effect.

$$E\left[Y_{ui}\,\text{¨O}do\left(T_{ui}=1\right)\right] \tag{2}$$

The deconfounding method produces superior recommendation results with better fairness than matrix-factorization when analyzing big datasets such as MovieLens 20M and Netflix Prize according to research [12]. The method produces superior results when working with sparse data and when creating recommendations for users who have not interacted before compared to the baseline method. The research built upon previous studies through deep latent-variable models and causal regularization methods which improved user diversity modeling [9].

The deconfounded recommender system framework establishes vital principles for causal recommendation modeling which enables recommendation evaluation to transition from assessing relevance to assessing intervention effects.

2.2 Identifiable Latent Confounder Modeling

When the observable features X_ui exhibit sufficient diversity, the latent confounding variable Zbecomes identifiable.

This structure enables the recovery of latent confounding factors that can be interpreted as user intent or environmental context. Experimental results on the Amazon Review dataset demonstrate that ILACs outperform black-box VAEs in debiasing performance while maintaining strong interpretability [9].

Moreover, identifiability ensures the consistency of causal estimation, which is essential for recommendation tasks in domains such as healthcare and education, where interpretability and reliability are critical. Integrating this framework with graph neural networks (GNNs) or disentangled representation learning can further enhance causal interpretability across multi-modal recommendation scenarios.

The process can be formalized as:

$$Z\rightarrow T_{ui}\rightarrow Y_{ui},Z\rightarrow X_{ui} \tag{3}$$

The observable features X_ui need to show enough diversity for the latent confounding variable Z to become identifiable.

The model structure allows researchers to detect vital concealed elements which represent user activities together with environmental conditions. The Amazon Review dataset experiments show that ILACs achieve better debiasing results than black-box VAEs while preserving excellent interpretability [7].

The method provides stable causal estimation results which recommendation systems need for their operation in healthcare and educational settings. The framework generates superior results for multi-modal recommendation tasks through its combination of graph neural networks (GNNs) with disentangled representation learning methods.

2.3 Causal Collaborative Filtering

The classical collaborative filtering method depends on user-item interactions to determine preferences yet fails to account for exposure mechanisms and contextual elements that affect results. The new modeling framework of Causal Collaborative Filtering (CCF) integrates causal reasoning into collaborative filtering to solve this problem [10]. The model distinguishes treatment effects from outcome effects through do(T_ui) prediction of Y_ui (T_ui) when the system takes action.

The model develops two separate latent factor spaces which represent exposure mechanisms and outcome production and calculates the causal effect through the difference between treated and control outcome predictions. The decomposition approach enables direct control of confounding variables while maintaining individualized recommendation capabilities.

The function can be expressed as:

$$\text{Preference}_{ui} = \underbrace{f_{\text{causal}}(u,i)}_{\text{treatmenteffect}} + \underbrace{?_{ui}}_{\text{confoundingnoise}} \quad (4)$$

The experimental results from KuaiRec and Yelp datasets show that CCF achieves superior model stability when distributions change and better fairness through decreased popular item exposure [10]. The model uses Pearl's do-calculus with gradient-based optimization to enable CCF for causal reasoning. The model learns from intervention-response relationships through end-to-end causal learning after implementing propensity weighting or latent confounder inference methods.

2.4 Causal Distillation and Uplift Modeling

Research on causal reasoning has developed into two separate methodological approaches which study human knowledge transfer and individualized strategy evaluation. The Causal Distillation method uses a teacher-student network structure where the causal teacher model learns domain-independent relationships which it then transfers to a simplified student model through causal consistency loss [8]. The model preserves stable "do-invariant" relationships which show how user characteristics affect outcomes at the intervention level making the model have cross domain generalization ability.

The causal teacher model trained on e-commerce data enables student model training for news recommendation to acquire fundamental causal principles which minimizes domain-specific influences. The method produces fair recommendations because it reduces the performance differences that occur between different user groups [8].

The marketing field developed uplift modeling known as boost modeling to enter the field of recommender systems [4, 13]. The main goal of uplift modeling differs from traditional absolute user response prediction because it focuses on determining the individual causal effect which represents the additional value users receive from exposure to specific items.

$$\tau_{ui} = E\left[Y_{ui} \text{¨O} T_{ui} = 1\right] - E\left[Y_{ui} \text{¨O} T_{ui} = 0\right] \quad (5)$$

The system learns τ_{ui} to determine which users need specific recommendations for targeted interventions.

The deployment of uplift-based recommendation systems produces superior marketing results and enhanced user satisfaction together with equal recommendation value for all users [4]. The combination of causal distillation with uplift (boost) modeling represents the main advancement for causal recommender systems.

The system enables causal recommendation to operate with meta-learning and personalized policy optimization through a unified framework which establishes a solid technical bond between causal reasoning and deep transfer learning for developing scalable interpretable causal personalization systems.

3. Experimental Datasets, Evaluation, and Applications

3.1 Datasets and Data Processing

The development of causal recommendation systems requires large observation datasets which need to include user feedback information and exposure tracking data.

The benchmark datasets for research include MovieLens and Netflix and Amazon Reviews and Yelp and KuaiRec but these datasets differ in their data sparsity levels and contextual diversity.The process of data preprocessing for causal inference needs more steps than standard collaborative filtering because it requires complete management of exposure mechanisms [1, 3, 12]. The following four aspects need to be addressed during data preprocessing.

3.1.1 Exposure Reconstruction

Most available datasets lack direct exposure logs so researchers use item popularity and ranking position and time decay indicators to create exposure probability estimates for items.

The propensity score estimation network [3] enables direct prediction of user item exposure likelihood.

Propensity Score Estimation and Weighting

The following formula determines the propensity score for each observation in the sample:

Assign a propensity score to each observation sample as follows:

$$\hat{p}_{ui} = P\left(T_{ui} = 1 \text{¨O} X_{ui}\right) \quad (6)$$

The IPS and DR estimators use this propensity score to remove biases from the data

Feature Extraction and Decontamination

The backdoor criterion in causal inference becomes possible through probabilistic matrix factorization and variational autoencoders (VAEs) which identify potential

ISSN 2959-6157

confounding factors to generate new features for dataset growth [14].

3.1.2 Construction of Counterfactual Datasets

The evaluation of untested recommendation approaches becomes possible through pseudo-intervention scenarios which either hide current exposure data or adjust sample combinations [10].

The analysis of long-term intervention effects on user engagement through repeated item exposure becomes possible through causal temporal graphs which handle dynamic confounding factors and time-dependent relationships in sequential data [5]. Healthcare and educational research need additional preprocessing techniques to manage confounding variables and missing data and censored results because these fields require strong causal inference methods [14].

The main goal of causal data preprocessing maintains observational data complexity while making causal relationships identifiable to preserve data diversity and maintain confounding factor traceability.

3.2 Evaluation Metrics

The assessment of causal recommendation systems requires evaluation methods that go past the conventional accuracy assessment techniques. The assessment of causal recommendation systems evaluates their unbiased performance and stable operation and fair treatment of users instead of using precision and recall and normalized depreciation cumulative gain (NDCG) and mean precision (MAP) metrics.

The Policy Value Estimation (PVE) system determines target policy returns through counterfactual learning by using three estimators which include inverse bias scoring (IPS) and self-normalized IPS (SNIPS) and dual robustness (DR) [15].

Treatment-Effect Metrics: The average treatment effect (ATE) and individual treatment effect (ITE) quantify the actual outcomes of recommendations which show the actual results of disambiguation.

Bias Reduction Indicators: The degree of distribution change determines the amount of bias reduction which affects selection bias and popularity bias and exposure bias [1].

The model stability assessment uses PVE and ATE variance measurements to evaluate its resistance to changes in subgroups and time segments [9]. The Gini index together with coverage rate and disparate-impact ratio enable researchers to assess both fairness and diversity in recommendation results [1].

The evaluation of causal explainability scores shows how well learned causal variables match explainable features to create more transparent algorithmic choices [9]. The evaluation process requires assessment of both predictive accuracy and maintenance of causal relationships in the model. The assessment method based on counterfactual risk minimization (CRM) analyzes recorded logged feedback during policy mismatches to learn from this data [16].

3.3 Application Scenarios and Empirical Findings

Causal recommendation systems have evolved from theoretical research to practical implementation across different domains.

The following five major application directions exist for causal recommendation systems. Dynamic and Sequential Recommendation: Temporal causal discovery technology enables effective modeling of feedback loops which enables the identification of short-term and long-term causal effects and reduces self-reinforcing popularity bias [5].

Healthcare Decision Support: The combination of causal reasoning with counterfactual prediction enables healthcare professionals to create treatment plans and predict patient outcomes which enhances decision accuracy [14].

The model enhances user targeting and personalized marketing results through its recommendation effect-based approach instead of using conversion rates [4].

Fairness-Aware E-commerce: The reweighting of product exposure through propensity estimation in causal recommendation systems produces equal product visibility and multiple recommendation options [1, 12].

Cross-Domain and Multi-Modal Recommendation: Causal Distillation allows the transfer of invariant causal knowledge between domains which produces superior results for cold-start items and multi-modal recommendations [8].

Research studies demonstrate that causal methods produce better results than traditional correlation-based methods because they enhance model stability and interpretability and fairness. Causal models transform recommender systems from predictive accuracy focus to decision-making accountability which results in more transparent and socially responsible systems.

4. Discussion and Future Perspectives

4.1 Current Challenges

Multiple obstacles exist when applying causal inference techniques to recommender systems because they affect both system readiness and theoretical advancement.

The identification of all possible confounding factors during the process becomes extremely difficult. The current practice of using proxy variables to estimate unobserved factors faces challenges because user behavior responds to changing preferences and multiple environmental factors which proxy variables struggle to represent correctly. Sequential recommendation systems experience dynamic processes because their recommendation strategies interact with user behavior through feedback loops which static causal models fail to model effectively.

The strict requirements of causal methods create an opposition with the operational requirements of system efficiency. The implementation of most causal models requires additional estimation and solution modules which result in higher training and inference expenses that prevent their use in real-time large-scale systems. The evaluation process encounters various challenges because researchers cannot determine cause-and-effect relationships and their offline assessment methods produce wrong results and online testing needs significant financial investment. The absence of standardized causal evaluation criteria prevents researchers from performing reliable comparisons between their studies.

The current combination of deep learning with causal inference methods faces technical barriers as well as moral challenges. Deep learning models excel at data representation yet their unexplainable operation methods create a conflict with the requirement for interpretability in causal analysis. The two methods need to find a balance which upholds both organizational structure and creative artistic elements. The identification of biases through causal models does not solve the problem of fairness because fairness exists in various forms which depend on particular situations. The theoretical framework needs additional development to establish common fundamental principles which will link recommendation systems with causal and counterfactual reasoning [17].

4.2 Future Directions

The future causal recommendation systems will unite neural networks with causality principles to create a complete system. The fundamental principles of causal reasoning will become part of neural network representation learning processes. The combination of causal graphical neural networks and causal variational autoencoders enables models to predict items more accurately while uncovering the causal relationships between recommended items.

The first phase of causal recommendation system development will bring revolutionary changes to the entire field of personalized recommendations. The system will advance from its current basic relevance prediction function to perform active causal reasoning which generates explainable and equitable results. The development of intelligent systems that understand and influence human decisions will achieve a major breakthrough through the successful execution of these innovative approaches.

5. Conclusion

The research provides an extensive analysis of causal reasoning in recommender systems which demonstrates how causal thinking enables new approaches to personalized recommendation development. Traditional recommender systems use models that predict user preferences but their correlation-based learning method creates selection bias and popularity bias and exposure bias which restricts their ability to produce unbiased results and maintain universal applicability. The theoretical framework of causal reasoning allows researchers to analyze user behavior mechanisms by using counterfactual analysis which generates both preference predictions and explanations of causal relationships.

Multiple causal methods have been developed as representatives during the last few years. The addition of proxy confounders in deconfounding recommenders helps remove hidden biases while models that detect latent confounders produce more consistent and easier-to-understand causal structures. Researchers developed counterfactual learning and evaluation methods which allowed them to build unbiased offline policy assessment systems that evolved into a counterfactual risk minimization framework which optimizes recommendation policies using log data. The „intervention-outcome" reasoning logic from causal collaborative filtering enables user-item interaction modeling and the causal distillation and enhancement models function together for cross-domain knowledge transfer and individualized effect estimation.

The new methods prove that causal inference methods enhance the fairness and robustness and interpretability of recommender systems. The development of causal recommender systems encounters three primary obstacles which require confounding factor detection and scalable causal estimator implementation and deep learning framework integration. Researchers need to use different data collection methods for causal model evaluation because they cannot measure actual causal outcomes directly. The solution to current problems requires researchers and developers to create methods for causal theory representation learning and standardized evaluation and fairness assessment protocols.

The theoretical framework of causal reasoning allows developers to build upcoming recommender systems which will advance from basic predictive tools to explainable decision systems that produce equitable and socially beneficial outcomes. The development of causal recommender systems will result in a single framework which unites

ISSN 2959-6157

causal logic with deep learning to execute reasoning and explanation and intervention capabilities. The practical applications of time-dimensional causal analysis and fairness-related learning and cross-domain knowledge transfer technologies will expand through the implementation of these technologies.

References

[1] Zhang S, Yao L, Sun A, & Tay Y. Deep learning based recommender system: A survey and new perspectives. ACM Computing Surveys (CSUR), 52(1), 1–38. 2019.

[2] Wang Y, Liang D, Charlin L, & Blei D M. The deconfounded recommender: A causal inference approach to recommendation. arXiv preprint arXiv:1808.06581. 2018.

[3] Peters J, Janzing D, & Schölkopf B. Elements of causal inference: Foundations and learning algorithms. The MIT Press. 2017.

[4] Gao C, Zheng Y, Wang W, Feng F, He X, & Li Y. Causal inference in recommender systems: A survey and future directions. ACM Transactions on Information Systems, 42(4), 1–32. 2024.

[5] Liu Z, Fang Y, & Wu M. Estimating propensity for causality-based recommendation without exposure data. Advances in Neural Information Processing Systems, 36, 51688–51705. 2023.

[6] Luo H, Zhuang F, Xie R, Zhu H, Wang D, An Z, & Xu Y. A survey on causal inference for recommendation. The Innovation, 5(2). 2024.

[7] Runge J, Gerhardus A, Varando G, Eyring V, & Camps-Valls G. Causal inference for time series. Nature Reviews Earth & Environment, 4(7), 487–505. 2023.

[8] Zhang S, Jiang Z, Yao J, Feng F, Kuang K, Zhao Z., ... & Wu F. Causal distillation for alleviating performance heterogeneity in recommender systems. IEEE Transactions on Knowledge and Data Engineering, 36(2), 459–474. 2023.

[9] Zhang Q, Zhang X, Liu Y, Wang H, Gao M, Zhang J, & Guo R. Debiasing recommendation by learning identifiable latent confounders. In Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (pp. 3353–3363). 2023.

[10] Xu S, Ge Y, Li Y, Fu Z, Chen X, & Zhang Y. Causal collaborative filtering. In Proceedings of the 2023 ACM SIGIR International Conference on Theory of Information Retrieval (pp. 235–245). 2023.

[11] Zhu Y, Ma J, & Li J. Causal inference in recommender systems: A survey of strategies for bias mitigation, explanation, and generalization. arXiv preprint arXiv:2301.00910. 2023.

[12] Guo R, Cheng L, Li J, Hahn P R, & Liu H. A survey of learning causality with data: Problems and methods. ACM Computing Surveys (CSUR), 53(4), 1–37. 2020.

[13] Moraes F, Manuel Proença H, Kornilova A, Albert J, & Goldenberg D. Uplift modeling: From causal inference to personalization. In Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (pp. 5212–5215). 2023.

[14] Prosperi M, Guo Y, Sperrin M, Koopman J S, Min J S, He X, ... & Bian J. Causal inference and counterfactual prediction in machine learning for actionable healthcare. Nature Machine Intelligence, 2(7), 369–375. 2020.

[15] Saito Y, & Joachims T. Counterfactual learning and evaluation for recommender systems: Foundations, implementations, and recent advances. In Proceedings of the 15th ACM Conference on Recommender Systems (pp. 828–830). 2021.

[16] Swaminathan A, & Joachims T. Counterfactual risk minimization: Learning from logged bandit feedback. In International Conference on Machine Learning (pp. 814–823). PMLR. 2015.

[17] Sharma A, & Kiciman E. Causal inference and counterfactual reasoning. In Proceedings of the 7th ACM IKDD CoDS and 25th COMAD (pp. 369–370). 2020.