

# 基于因果推断的推荐系统去偏研究综述

杨新新 刘真 卢思博 袁亚凡 孙永奇

(北京交通大学计算机科学与技术学院 北京 100044)

**摘要** 推荐系统是解决信息过载的重要技术之一,然而,推荐系统中存在各种各样的偏差问题,影响了对用户真实偏好的建模,制约了推荐性能的提升.近年来,因果推断理论的发展为分析和解决推荐系统偏差问题提供了有力的支持.因果推断是一种从观测数据中识别变量之间因果关系和估计因果效应的统计学方法,通过构建和分析因果模型,帮助推荐系统识别和消除偏差,提高对用户偏好拟合的准确性.本文对基于因果推断的推荐系统去偏研究的主要工作进行了全面的综述.本文根据推荐系统的不同阶段将偏差的产生分为三个阶段;首先,概述了因果推断的原理和方法,并阐述了因果推断与推荐去偏之间的联系,为缓解偏差问题提供了思路;接着,针对每一阶段的偏差,探讨了现有的因果推断技术如何应用于推荐去偏,分类和归纳了现有的因果推荐去偏方法,并进行了详细的对比分析;最后,讨论和展望了基于因果推断的推荐系统去偏研究未来的发展趋势.

**关键词** 推荐系统去偏;因果推断;偏差;反事实推理

中图法分类号 TP311 DOI号 10.11897/SP.J.1016.2024.02307

## A Survey on Debiasing Recommendation Based on Causal Inference

YANG Xin-Xin LIU Zhen LU Si-Bo YUAN Ya-Fan SUN Yong-Qi

(School of Computer Science and Technology, Beijing Jiaotong University, Beijing 100044)

**Abstract** Recommender systems play a vital role in addressing information overload by learning user preferences from historical interaction data and thereby providing personalized recommendations. However, various biases in recommendation systems hinder the accurate modeling of users' true preferences, limiting the improvement of recommendation performance. Recently, the development of causal inference theory has provided robust support for analyzing and resolving bias problems in recommender systems. Causal inference, a statistical method aimed at identifying and estimating causal effects between variables from observational data, assists in identifying and eliminating biases through the construction and analysis of causal models, enhancing the accuracy of fitting user preferences. Applying causal inference to debiasing tasks in recommender systems has achieved significant success, effectively mitigating bias while also enhancing accuracy and reliability. This paper provides a comprehensive review of the research on debiasing recommendations based on causal inference. Considering that bias occurs at various stages of recommender systems, we classify the sources of bias according to the three stages of recommender systems: data, algorithms, and evaluation. We also summarize the manifestations of bias at each stage and their impact on recommendations. Based on the study of debiasing recommendations from a causal perspective, we first outline the principles and key methods of causal inference. This establishes the connection between causal inference and debiasing recommendation, providing insights into mitigating bias.

收稿日期:2023-11-18;在线发布日期:2024-07-08.本课题得到国家重点研发计划项目(2019YFB2102500)资助.杨新新,博士研究生,主要研究方向为推荐系统、因果推断. E-mail: yangxinxin@bjtu.edu.cn.刘真(通信作者),博士,教授,博士生导师,中国计算机学会(CCF)高级会员,主要研究领域为数据挖掘、推荐系统、分布式计算. E-mail: zhliu@bjtu.edu.cn.卢思博,硕士研究生,主要研究方向为推荐系统、因果推断.袁亚凡,硕士研究生,主要研究方向为推荐系统、因果推断.孙永奇,博士,教授,博士生导师,中国计算机学会(CCF)会员,主要研究领域为数据挖掘、人工智能.

Then we systematically organize and analyze debiasing strategies for recommender systems at the data, algorithm, and evaluation stages based on causal inference techniques. For debiasing methods at the data stage, there are primarily two strategies based on how the data is utilized: counterfactual construction-based methods generate synthetic data points to simulate what might happen under different scenarios, helping to uncover hidden biases; unbiased data-based methods involve collecting data that is free from the common biases present in real-world datasets. For debiasing methods at the algorithm stage, there are primarily three strategies based on different causal techniques: causal representation learning-based methods aim to learn representations of the data that are invariant to biases; causal intervention-based methods directly manipulate variables to observe changes and infer causal relationships; counterfactual reasoning-based methods involve comparing actual outcomes with hypothetical scenarios to identify and correct biases. For debiasing methods at the evaluation stage, there are primarily two strategies based on the correction and optimization of unbiased estimates: inverse propensity scoring-based methods adjust for the probability of receiving a particular treatment, helping to balance the dataset; doubly robust-based methods combine propensity score weighting with outcome modeling to improve the robustness and accuracy of bias correction. Currently, recommender systems based on causal inference represent a novel and challenging research field. This paper summarizes several open research directions, including debiasing recommendation methods based on causal discovery, a general causal-based debiasing recommendation framework, robust debiasing methods based on causal inference, addressing bias issues in dynamic environments using causal approaches, and the construction of datasets for causal debiasing recommendations. Finally, we summarize this paper and provide an outlook on the research of debiasing recommendations based on causal inference from the perspectives of application needs and technological development.

**Keywords** debiased recommendation; causal inference; bias; counterfactual inference

## 1 引 言

在这个信息过载的时代,作为一种信息过滤技术,推荐系统<sup>[1-2]</sup>取得了巨大的成功.推荐系统通过从用户的历史交互数据中学习用户的偏好,从而为用户提供个性化的推荐服务.不管是传统的基于协同过滤的方法,还是近年来基于深度学习的推荐方法,都致力于设计更好的模型拟合用户的历史数据来提升推荐质量.目前,推荐系统不仅是学术界的研究热点,取得了众多的研究成果,而且在工业界得到了广泛的应用,如阿里巴巴<sup>[3]</sup>、谷歌<sup>[4]</sup>和 Youtube<sup>[5]</sup>等.总之,推荐系统在当今的互联网时代扮演着极为重要的角色.

然而,推荐系统在实践中面临着偏差问题<sup>[6-9]</sup>.一方面,由于用户的历史交互数据存在缺失和不平衡的情况,使得推荐数据先天存在各种各样的偏差,比如选择偏差、曝光偏差、位置偏差等.偏差的存在使得训练数据分布偏离用户真实兴趣分布,影响了

推荐的准确性.另一方面,推荐算法自身在表征学习过程中同样会产生大量的偏差,它们在推荐过程中会倾向于向交互频繁的用户推荐头部商品或热门商品,而忽略了其他用户和商品,导致不准确、不全面的推荐结果<sup>[10]</sup>.除此之外,推荐系统在评估过程中通常依赖历史数据,可能无法准确反映当前的用户偏好和行为模式,同时评估指标和评估方法也可能引入偏差,从而影响推荐系统的性能和用户满意度.因此,解决推荐系统的偏差问题成为了推荐系统领域的一个重要研究方向,并获得了广泛的关注.传统的推荐系统去偏方法主要通过补充相似的数据源来处理数据偏差,但是传统的去偏方法存在很多限制,如无法处理缺失数据和噪声、缺乏鲁棒性和可解释性等.随着因果推断技术的进步,研究者们开始将基于因果的方法应用到推荐系统去偏的研究上,这些基于因果的推荐方法克服了传统方法的不足,增强了模型的准确性和稳定性.

因果推断(Causal Inference)是一种统计学方法,旨在更科学地研究变量之间的因果关系<sup>[11-12]</sup>.由

于其拥有更庞大的可用数据和更低的预算要求,与随机对照实验相比,从观察数据中估计变量的因果效应成长为一个极具潜力和吸引力的研究方向。同时,因果推断与机器学习领域有着密切的联系<sup>[13-14]</sup>。一方面,机器学习促进了因果推断领域的发展,强大的机器学习方法可以更准确地估计因果效应。另一方面,因果推断也有助于机器学习方法的发展,从因果的角度提高现代机器学习方法的准确性和解释性。而推荐系统与因果推断也是高度相关的<sup>[15-16]</sup>。推荐系统的本质是建模用户的偏好并预测用户的下一次交互,而这一目标可以被合理地抽象为一个具体的因果推断问题。举例来说,在电影推荐场景下,预测的过程其实是推荐系统回答“用户是否会观看这部电影?”这样一个关于因果推断的问题。在因果推断框架下,推荐行为作为一种“干预”,旨在探究如果推荐某部电影给用户,用户是否会因此选择观看,相比之下,不进行推荐时用户的行为可能有所不同。这样的干预不仅考量了推荐系统的潜在增益,更重要的是关注如何更好地理解和满足用户的真实需求和兴趣。在这一例子中,系统“推荐”是因,用户“观看”是果,更详细的信息在第3节会进一步介绍。

最近,许多工作开始将因果推断技术应用到推荐系统领域,尤其在处理推荐系统偏差问题上取得了显著的成效。实践表明,基于因果推断的推荐去偏方法有以下特点和优势:首先,传统的推荐方法通过相关性学习来拟合数据,无法处理观察数据里的缺失或混杂因素,使得预测结果偏离用户真实的喜好。而因果推断通过建立因果模型来推断变量之间的因果关系,从而帮助厘清造成偏差的原因,实现更可靠的推荐。其次,对于数据中可观察到的偏差,因果推

断可以通过干预的手段更好地控制偏差的影响,提高推荐的准确性。最后,因果推断通过因果分析可以提高推荐系统的可解释性,帮助用户理解推荐结果背后的原因,增加对推荐系统的信任。综上所述,基于因果推断的推荐方法不仅能够有效缓解偏差问题,具有更高的准确性和可靠性,还能够提供可解释性。

随着推荐系统去偏和因果推断研究的发展,近几年逐渐有相关研究综述发表<sup>[6,17-18]</sup>。其中,文献[6]首次对推荐系统领域的偏差问题和去偏方法进行了统一的定义和总结,文献[17]系统地论述了如何使用因果增强推荐系统。这些综述侧重于对推荐系统偏差的描述或因果推断在整个推荐领域的应用,但是缺乏对因果去偏任务的系统论述。本文聚焦于基于因果推断的推荐系统去偏研究,旨在提供针对这一具体领域的最新进展综述,弥补研究空缺。本文从因果推断技术的角度出发,针对推荐系统的偏差及去偏问题,对相关领域具有前沿性和代表性的文献进行归纳总结,系统地概述了因果推断在推荐系统去偏任务中的应用,并对现有方法的原理和优缺点进行了分析。最后展望了基于因果推断的推荐系统去偏研究的未来趋势。本文第2节详细说明推荐系统存在的偏差问题并根据偏差的来源对其进行分类;第3节介绍因果推断的背景知识,建立起因果推断与推荐系统去偏之间的联系;第4~6节根据推荐系统中偏差产生的不同阶段——数据、算法和评估,分类讨论各阶段的去偏方法,并从因果推断的角度对去偏方法进行分类和分析;第7节讨论基于因果推断的推荐系统方向中一些开放性的研究问题;第8节对本文进行总结。本文的整体架构如图1所示。

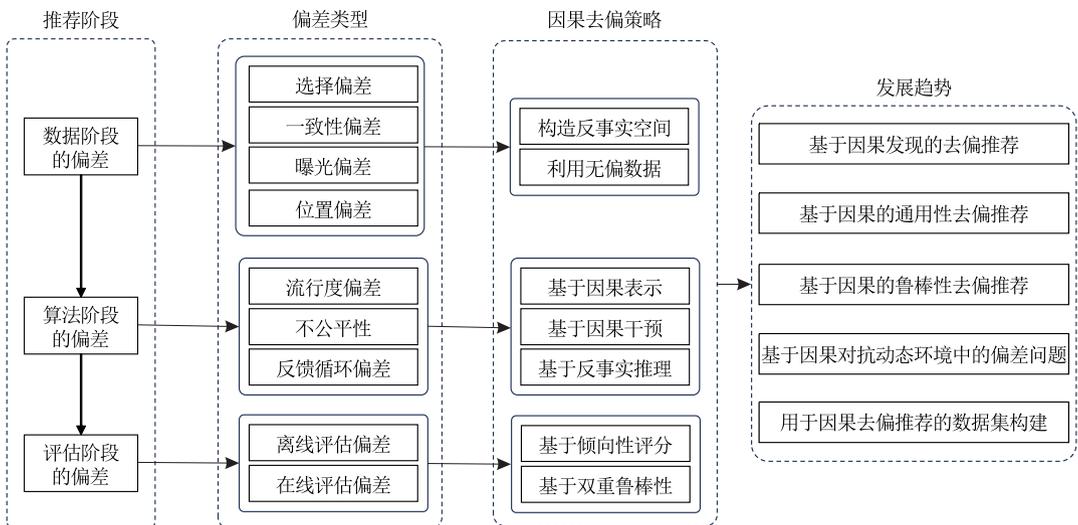


图1 本文的整体架构

## 2 推荐系统的偏差问题及挑战

偏差一般指训练模型的预测结果与真实值之间的差异。推荐系统在为用户提供推荐时,可能会因为某些因素导致推荐结果与用户的真实偏好产生偏差,从而降低推荐系统的质量和用户的满意度。不同形式的偏差可能在各个阶段影响推荐系统。一个完整的推荐模型构建过程可以划分为数据、算法和评估三个阶段:数据处理是推荐系统的重要基础,主要任务是处理原始数据,进行特征提取和转换;算法阶段通过选择、建立和训练合适的模型来进行预测和推荐;最后通过评估推荐系统的性能,对数据处理和算法进行调整和优化<sup>[19]</sup>。以上三个阶段是推荐系统的核心环节,也涵盖了推荐系统构建的完整过程。推荐系统在这三个阶段都可能产生偏差,因此本文按照这三个阶段对推荐系统中的偏差进行了分类,即数据阶段的偏差、算法阶段的偏差和评估阶段的偏差,如图 1 所示。考虑到各阶段也可能产生相互影响,本文对某些受到多个阶段影响的偏差会进行具体的分析。下文将展开说明图 1 中各偏差的来源以及它们如何对各阶段产生影响。

### 2.1 数据阶段的偏差

数据阶段的偏差指推荐系统在为用户提供服务时,所使用的用户历史行为数据中存在的偏差。数据偏差源于用户互动和系统展示的非随机性,导致收集的用户历史行为数据不完全代表所有用户真实的偏好。数据阶段偏差主要包含以下几种:选择偏差、一致性偏差、曝光偏差和位置偏差等。这些数据偏差可以根据偏差产生的主体分为两类:选择偏差和一致性偏差是由用户行为导致的偏差,曝光偏差和位置偏差是由推荐系统对物品的展示导致的偏差。

#### 2.1.1 用户行为导致的偏差

选择偏差<sup>[20-21]</sup>指观察到的用户评分数据是用户根据自己的偏好自主选择评分的,导致观测到的数据是一种非随机缺失性(Missing Not At Random, MNAR)数据。这种数据不能全面代表所有物品的潜在评价,因为用户倾向于只与他们预先感兴趣的物品互动。例如,在电影评分平台,用户倾向于对自己喜欢的电影进行评分,却很少对他们不喜欢的电影进行评分。选择偏差指的就是用户自主评分的数据与随机分配物品给用户评分的数据分布不一致的现象。

一致性偏差<sup>[22-23]</sup>指用户在评分时容易受到他人

评价的影响,以至于该用户的评分不能直接反映其真实兴趣。例如,用户往往会受到朋友的影响而给出和朋友相似的评分。还有些用户看到大多数人对某个物品打高分或低分时,也会出于从众心理选择和大多数人的评价趋同,这是一种很正常的社会现象。因此观察到的评分数据往往是有偏的。

#### 2.1.2 推荐系统导致的偏差

曝光偏差<sup>[24-25]</sup>指推荐系统由于算法或设计选择,仅将部分特定的物品曝光给了用户,而忽略了其他可能同样符合用户兴趣的物品,导致未被系统曝光的物品缺乏用户互动记录。曝光偏差主要存在于隐式反馈数据中,即用户的点击、浏览或收藏等行为。这些行为只能反映用户对曝光过的物品的喜好,而不能反映那些未曝光的物品。在曝光偏差的影响下,推荐系统倾向于反复推荐已被曝光过的物品,而放弃探索新的或未曝光的物品。

位置偏差<sup>[26-27]</sup>指用户倾向于与推荐列表顶部的项目进行交互,而忽略一些相关性较高却位于推荐列表底部的物品。位置偏差通常存在于基于排序的推荐系统中,比如数字图书馆、搜索引擎或电子商务等领域。在这些系统中,用户往往只浏览或者点击位置靠前或醒目的一些项目,导致这些项目随着时间的推移被过分推荐,从而影响推荐的准确性和多样性。

### 2.2 算法阶段的偏差

算法阶段的偏差指推荐算法本身存在的偏差,即算法的设计或实现导致的对某些物品或用户的不公平或不合理的处理。主要包含以下几种:流行度偏差、不公平性偏差和反馈循环偏差。

流行度偏差<sup>[9,22]</sup>指推荐算法倾向于给用户推荐更为流行的物品,从而减少了流行度低但可能更受用户喜爱的物品被推荐的机会。流行度偏差来源于以下原因:由于用户的行为数据本身存在长尾分布,即训练数据中大多数的用户-物品交互记录被极少数热门物品所占据。在长尾分布数据上训练的推荐算法倾向于根据物品的流行度来推荐物品,因为从模型优化的角度分析,更频繁地推荐热门物品有利于实现更小的损失,使得模型从该方向更新参数。因此,流行度偏差阻碍了推荐系统对用户兴趣的理解,并且损害部分用户的体验。同时,随着推荐系统的循环,流行度偏差会进一步加剧,导致马太效应,降低推荐系统的个性化。

不公平性偏差<sup>[8,23]</sup>推荐算法对特定的群体有所偏袒,而对某些群体或个人有歧视,比如性别歧视、

种族歧视或年龄歧视等。不公平性主要来源于以下原因：首先，由于推荐系统中的用户数据通常是不平衡的，因此某些用户可能会被低估甚至被忽略，这也是推荐领域一直以来的难点之一。其次，不同的推荐算法本身也会带来一定的偏差，比如基于流行度的算法会更加偏爱热门的物品，基于协同过滤的算法则会强化用户和物品之间的相似性，在某种程度上忽略用户的个性化需求。不公平性偏差除了来自算法偏差之外还受到数据偏差的影响。在推荐系统中，不公平性会降低推荐的准确性和多样性，严重甚至会引发社会问题。

反馈循环偏差<sup>[10,28]</sup>指在推荐系统循环过程中偏差被逐步放大的问题。文献[8]将推荐系统总结为“用户-数据-模型-用户”这样一个反馈回路。在此反馈回路中，用户和推荐系统处于相互动态演变的过程中，用户的行为会随着时间的推移通过推荐系统而更新，推荐系统的训练也会受到新的用户行为数据影响。反馈循环偏差指的是在这一动态过程中，源自用户历史交互数据里的偏差会导致推荐模型的结果存在偏差，推荐模型导致的偏差反过来会进一步加剧交互数据里的偏差。如此循环，偏差会被一步步放大，从而降低推荐的多样性和用户的同质性，导致马太效应。因此，反馈循环偏差同时存在于算法阶段和数据阶段。

### 2.3 评估阶段的偏差

评估阶段的偏差指对推荐系统性能进行评估时，所使用的评估指标、数据集或评估方法存在一定

的差异，不能准确地反映用户的满意度或商业指标。主要包含以下几种：离线评估偏差和在线评估偏差。

离线评估偏差<sup>[29-30]</sup>指在离线评估过程中使用的数据集和评估指标与实际用户行为之间存在差异，导致离线评估结果与在线评估结果不一致。离线评估偏差主要来源于以下几种原因：由于用户历史行为数据本身存在的数据偏差导致训练数据与能够代表用户真实兴趣的数据分布不一致，影响了模型的评估；由于用户的兴趣和需求是动态变化的，导致数据集中的用户反馈可能已经过时或不适用于当前的场景，影响了推荐模型的泛化能力和适应性；由于推荐模型的效果涉及多个方面，如准确性、新颖性、多样性、覆盖率等，导致单一的评估指标或方法可能不能全面地衡量推荐模型的优劣。

在线评估偏差<sup>[31-32]</sup>指在在线环境下，使用真实用户流量对推荐模型进行评估时存在的不足。在线评估偏差主要来源于以下几种原因：由于用户在真实场景下进行交互时，可能受到其他用户、社会、环境等外部因素的影响，导致用户的反馈不能完全反映用户对推荐模型的满意度；由于用户在实际场景下进行交互时，可能受到推荐模型本身的影响，如过度个性化、过度泛化、过度拟合等，导致用户的反馈不能完全反映用户对推荐物品的兴趣。

表1总结了推荐系统中主要偏差的来源、分类和偏差表现。除了上述常见的偏差之外，还有一些数据偏差如点击诱饵偏差<sup>[33]</sup>、情感偏差<sup>[34]</sup>和点击偏差<sup>[35]</sup>等逐渐开始得到研究者们的关注。

表1 偏差的分类与分析

偏差来源	偏差类别	偏差表现	相关文献
数据偏差	选择偏差	观察到的评分分布与实际评分分布不一致	[20, 21]
	一致性偏差	用户因从众心理行为受到他人影响	[22, 23]
	曝光偏差	少数被曝光的物品无法反映用户真实的偏好	[24, 25]
	位置偏差	推荐列表头部的更容易被交互	[26, 27]
算法偏差	流行度偏差	热门商品被过度推荐	[9, 22]
	不公平性	对少数群体或个人有歧视	[8, 23]
	反馈循环偏差	偏差经过推荐系统反馈回路被放大	[10, 28]
评估偏差	离线评估偏差	评估数据和指标与实际用户行为不一致	[29, 30]
	在线评估偏差	在线评估不能完全反映真实效果	[31, 32]

## 3 因果推断概述

因果推断是一种统计方法，专注于研究和推断因果关系。在推荐系统领域，因果推断被广泛用于解决推荐系统的偏差问题。因果推断为推荐系统提供了一种强大的工具，使得能够通过审查推荐结果的生成过程来确定偏差的根本原因，并通过因果推荐

建模减轻偏差的影响，提高推荐结果的准确性和公平性。为了更好地从因果推断技术的角度出发，对因果去偏推荐方法进行综述，本节将介绍因果推断的背景和原理，以及与推荐系统去偏的关联。

贝叶斯网络之父 Pearl<sup>[36]</sup>将通向强人工智能的因果学习划分为三个层次：关联(Association)、干预(Intervention)以及反事实(Counterfactual)。第一层的“关联”是指基于观察到的数据样本来学习变量

之间的关联关系,未考虑因果关系,其主要目标是实现预测;第二层的“干预”是指有计划地采取行动,主动获取新的知识,探索因果关系,目的是做出更优的决策;第三层的“反事实”是指利用反事实推理来探索未观测到的数据,扩展想象空间,摆脱现实世界的束缚,创造新的事物。

下文首先介绍两种基本的因果框架:结构因果模型和潜在结果框架,然后进一步描述重要的因果推断方法和原理,包括逆倾向性得分方法、双重鲁棒性方法、因果干预方法和反事实推理方法。

### 3.1 因果推断框架

#### 3.1.1 结构因果模型

结构因果模型 (Structural Causal Models, SCM)<sup>[37]</sup> 是一种基于图论的因果框架。结构因果模型通过建立变量之间的有向无环图 (Directed Acyclic Graph, DAG) 来表示因果关系,然后利用干预的手段来进行因果推理,从而挖掘数据产生的机制和因果关系。从形式上看,结构因果模型由两部分组成:因果图和对应的一组结构方程。因果图上的节点表示变量,边表示变量之间的因果关系。在因果图中,如果变量  $B$  是变量  $A$  的子节点,那么变量  $A$  是变量  $B$  的直接原因。结构方程用于建模因果图中的因果关系,通过对应的结构方程,因果图中的每一个变量可以根据其父节点计算得出。

根据有向无环图的路径结构,结构因果模型中的基础结构可以分为以下三种:链状 (chain) 结构、叉状 (fork) 结构和对撞 (collider) 结构,如图 2 所示。图 2 中,链状结构为  $X \rightarrow Y \rightarrow Z$ ,表示信息仅可单向流通;叉状结构为  $X \leftarrow Y \rightarrow Z$ ,表示信息可以从中间分发到两端;对撞结构为  $X \rightarrow Y \leftarrow Z$ ,表示中间同时接收两端节点的信息。在结构因果模型中不管多么复杂的结构都可以拆分为以上三种结构的组合,这三种结构也可能导致不同的偏差。在复杂因果模型的拆解分析中需要考虑全部因果路径,才能推断出准确的因果关系。举例来说,在推荐系统中,链状结构“兴趣  $\rightarrow$  点击”表示由用户兴趣引发的点击行为,具有因果性;叉状结构“价格  $\leftarrow$  质量  $\rightarrow$  点击”表示物品的质量同时影响物品的价格和用户的点击,如果忽略质量这一变量,物品的价格和用户的点击行为之间呈现伪相关;对撞结构“兴趣  $\rightarrow$  点击  $\leftarrow$  流行度”表示用户的兴趣和物品的流行度共同导致了用户对该物品的点击,如果控制点击这一变量不变,兴趣和流行度之间也呈现伪相关。

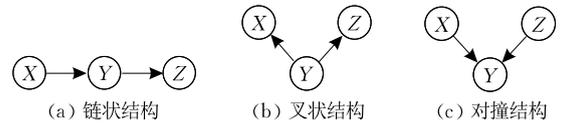


图 2 因果图结构

#### 3.1.2 潜在结果框架

潜在结果框架,也称为鲁宾因果模型 (Rubin Causal Model, RCM)<sup>[38]</sup>,是一种基于反事实逻辑的因果推断框架。潜在结果框架的基础是,因果关系与作用于某一单元 (unit, 如一个用户) 的干预 (treatment, 如推荐不同的商品) 有关,通过比较不同干预下的潜在结果可以估计因果效应。潜在结果是指给定一个单元和一系列干预,把一个“干预-单元”确定为一个潜在结果。例如,如果给用户  $A$  推荐商品  $X$ ,那么用户  $A$  对商品  $X$  的评分就是一个潜在结果。但是,在实际中,只能观察到用户  $A$  实际接受的那一种推荐的结果,而无法观察到其他可能存在的推荐下的结果,这些未观察到的推荐结果被称为反事实结果。因此,潜在结果框架的主要目标是估计不同干预下的潜在结果 (包括反事实结果),以估计实际的干预效果。

潜在结果框架的干预效应可以通过以下公式来描述。首先,从单元层面上,干预效应可以描述为平均干预效应 (ATE):

$$ATE = \mathbb{E}[Y(W=1) - Y(W=0)] \quad (1)$$

其中,  $Y(W=1)$  和  $Y(W=0)$  分别为整体的潜在干预结果和对照结果。

其次,在个体层面上,干预效应可以描述为个体干预效应 (ITE):

$$ITE_i = Y_i(W=1) - Y_i(W=0) \quad (2)$$

其中,  $Y_i(W=1)$  和  $Y_i(W=0)$  分别为对某一单元  $i$  的潜在干预结果和对照结果。

因果推断的目标是通过以上原理从已观测数据中估计干预效应。当然,上述因果推断的理论都需要满足非常强的假设才能成立。下面介绍三种最基础的假设。

(1) 稳定性假设 (SUTVA) 指任一单元的潜在结果不受其他单元的干预分配影响,也就是说,单元之间是相互独立的,单元之间不互相干扰。同时,对于每个单元的干预分配只有一种形式或版本,因为不同的干预形式或版本会导致不同的潜在结果。

(2) 可忽略性假设 (Ignorability) 指给定环境变量  $X$ ,对某一个体的干预分配机制  $W$  与潜在结果相互独立,即  $W \perp Y(W=0), Y(W=1) | X$ 。在可忽略

性假设下,具有相同环境变量下的单元对于这些单元的干预分配可以看作是随机的。

(3) 正值假设(Positivity)指对于任何  $X$  值,干预分配都是不确定的:  $P(W = w | X = x) > 0, \forall w$  and  $x$ . 换句话说,对于具有不同  $X$  值的任何单元组都有一定的概率接受任何一种干预。

相对于结构因果模型,潜在结果框架可以直接估计因果效应,不需要假设因果机制或建立因果图。但是潜在结果框架需要满足一些较强的假设,如上文提到的稳定性假设、可忽略性假设和正值假设。同时,潜在结果框架往往需要大量的数据或随机实验来消除混杂变量的影响。而结构因果模型则可以借助图模型进行更复杂的因果推理。然而结构因果模型需要假设因果图和结构方程,并且可能存在识别问题或参数估计问题等。而且,结构因果模型往往需要专业的知识或工具来构建和分析。以上两种模型在机器学习包括推荐系统领域都有广泛的应用,且两种模型已经被证明是等价的<sup>[39]</sup>。从应用角度看,潜在结果框架更加精确,而结构因果模型更加直观。

### 3.2 因果推断方法

#### 3.2.1 逆倾向性得分

倾向性得分(P propensity Score)<sup>[40]</sup>是指一个样本被分配到处理组的概率,它通常用观察到的协变量来估计。逆倾向性评分(Inverse Propensity Score, IPS)<sup>[41]</sup>是一种基于倾向分数的加权方法,使得处理组和对照组之间的协变量分布相似。它的基本思想是,如果一个样本被分配到处理组的概率很低,那么它在处理组中的权重就应该很高,反之亦然。IPS 通过计算每个个体接受处理的概率,然后将这个概率作为权重,对每个个体的结果进行加权,从而得到一个加权样本的平均结果。这个加权过程可以减少混淆因素的影响,从而更准确地推断出处理对结果的影响。具体来说,IPS 方法的步骤如下:(1) 估计每个个体接受处理的概率,即倾向分数;(2) 计算每个个体的 IPS 权重,即倾向分数的倒数,这个权重表示个体的重要性;(3) 根据 IPS 权重对每个个体的结果进行加权平均,得到处理组和对照组的平均结果;(4) 比较处理组和对照组的平均结果,推断出处理对结果的影响。基于 IPS 的无偏估计器可以用如下经验风险最小化函数表示:

$$\arg \min_{\theta} \sum_{O_{u,i}=1} \frac{\hat{\delta}_{u,i}(r, \hat{r}(\theta))}{P_{u,i}} + \text{Reg}(\theta) \quad (3)$$

其中,  $\hat{\delta}_{u,i}(r, \hat{r}(\theta))$  指原始的估计函数,  $\text{Reg}(\theta)$  是防

止过拟合的正则项,倾向分数  $P_{u,i}$  是作用于每一个损失项的权重。

在推荐系统中,数据偏差会导致观测到的用户-项目评分与用户潜在兴趣的真实分布不一致。为了缓解这一问题,可以基于 IPS 加权来修正损失函数,构造无偏的估计器。其中,倾向分数的应用可以通过优化无偏的经验风险函数来有效缓解选择偏差。本文在推荐系统评估阶段的去偏方法研究中展示了实际推荐系统中应用 IPS 的过程和效果,讨论了 IPS 去偏策略的优势和挑战。

#### 3.2.2 双重鲁棒性

双重鲁棒性(Doubly Robust, DR)<sup>[42-43]</sup>是一种基于模型的调整方法,它可以同时利用倾向性评分和回归模型来处理混淆因素的影响。DR 的基本思想是通过建立一个模型来估计处理和结果之间的关系,然后根据模型进行加权和调整,从而得到更准确的结果。与 IPS 方法不同的是,DR 方法不仅考虑了处理和结果之间的关系,还考虑了其他因素对结果的影响。双重鲁棒性的含义是,只要倾向得分或回归模型中的一个是正确的,那么最终的估计就是一致的。具体来说,DR 方法的步骤如下:(1) 建立一个模型来估计处理和结果之间的关系;(2) 估计模型的参数,得到一个对结果的预测值;(3) 估计每个个体接受处理的概率,即倾向分数;(4) 根据倾向分数和模型预测值对每个个体的结果进行加权平均,得到处理组和对照组的平均结果;(5) 比较处理组和对照组的平均结果,推断出处理对结果的影响。双重鲁棒性方法的目标函数如下:

$$\mathcal{E}_{\text{DR}} = \mathcal{E}_{\text{DR}}(\hat{r}, r^o) = \frac{1}{nm} \sum_{u,i} \left( \hat{e}_{u,i} + \frac{o_{u,i} d_{u,i}}{\hat{p}_{u,i}} \right) \quad (4)$$

其中,  $\hat{e}_{u,i} = \delta_{u,i}(r_o, \hat{r}_{u,i})$  表示预测值和插补值之间的估计误差,  $d_{u,i} = \hat{e}_{u,i} - \delta_{u,i}(r, \hat{r})$  表示插补误差和预测误差之间的偏差。 $n$  和  $m$  分别为用户和项目的数量。

相对于 IPS 方法,DR 方法的优点是它同时考虑了处理和结果之间的关系和其他因素对结果的影响,因此比 IPS 方法更具有鲁棒性。但它需要建立一个模型来估计处理和结果之间的关系,如果模型不准确,将会影响结果的准确性。

在推荐系统的去偏应用中,DR 估计器通过修正插补误差并结合倾向分数的逆加权来提高估计的准确性。当用户-项目之间存在历史交互记录时,DR 估计器通过计算预测误差和插补误差的加权组合来调整推荐结果,而在无历史交互的场景中,估计器专

注于插补误差的调整. 实验表明,即使在倾向性评分或回归模型中有一项存在偏差,DR 方法仍能维持其无偏估计的属性. 与传统的 IPS 方法相比,DR 方法在提高推荐鲁棒性方面有显著优势,但是在实践中需要谨慎选择合适的回归模型,并确保模型的准确性.

### 3.2.3 因果干预

上文提到过因果关系的三个层级:关联、干预和反事实. 其中关联是依据已观测到的数据进行统计和预测,分析变量之间的相关性,并不改变原有数据的分布. 这种基于观测数据的相关性研究无法厘清变量之间的因果关系,而干预则需要改变原有的数据分布. 在因果图模型中,干预是指通过修改某些变量的值来观察其他变量的变化,从而确定变量之间的因果关系. 具体地说,当对因果图模型里的某一个变量实施干预时,固定住该变量的值,这时该变量不会受其他变量的影响而变化. 那么在因果图模型里,干预就是将所有指向该变量的边删除,改变了图的结构和变量的相关性<sup>[36]</sup>.

$do$  表达式是一种用来计算干预效应的数学工具. 当固定变量  $Y$  的值为  $y$  时,用数学符号  $do(Y=y)$  来表示这种干预行为. 所以  $P(Z=z|Y=y)$  表示当以  $Y=y$  为条件时  $Z=z$  的概率,  $P(Z=z|do(Y=y))$  表示当干预  $Y$  的值使其为  $y$  时,  $Z=z$  的概率. 从概率分布的角度来说,  $P(Z=z|Y=y)$  表示的是在  $Y$  可取的所有值中,  $Y=y$  那部分样本对应的  $Z=z$  的概率,而  $P(Z=z|do(Y=y))$  表示的是将每一个样本的  $Y$  的值全部固定为  $y$  后  $Z=z$  的概率. 这两者是完全不同的,干预改变了原始数据的分布,而以变量为条件不改变原始数据的分布,使  $do$  表达式和对因果图的操作能够从关联关系中厘清因果关系.

接下来重要的是如何估算  $do$  表达式下的条件概率,即如何在概率计算的过程中消去  $do$  表达式. 前面提到如果对某变量进行干预,需要对该变量的父节点进行校正,但实际上变量往往会受到一些不可观测的父节点影响,此时需要找到一个替代变量集合来用于校正. 因此, Pearl 等人<sup>[37]</sup> 提出了一个重要的因果干预工具—后门准则来估计因果效应.

后门准则指出,在 DAG 中,若变量集合  $Z$  相对于一组变量有序对  $(X, Y)$  满足后门准则,则 (1)  $Z$  中节点不能是  $X$  的后代; (2)  $Z$  阻断了  $(X, Y)$  之间所有指向  $X$  的路径(这样的路径可以称为后门路径). 进一步说,如果存在一个变量集合  $X$  和  $Y$  满足后门

准则,那么变量  $X$  到  $Y$  的因果效应可以通过如下公式进行计算:

$$P(Y=y|do(X=x)) = \sum_z P(Y=y|X=x, Z=z)P(Z=z) \quad (5)$$

通过对变量集合  $Z$  做调整,即先分层再加权求和,从而消除  $do$  表达式. 至此可以得到  $do$  表达式下的条件概率来估计因果效应.

在推荐系统中,因果干预提供了一个强有力的工具,允许通过  $do$  表达式实施干预并估计其效果. 例如,为了克服流行度偏差的影响,可以设计实验对推荐算法进行调整,以消除流行度因素对用户兴趣预测的影响. 通过后门准则,可以识别并调整推荐系统中的混杂因素,如用户的历史行为模式,以确保预测更加准确. 通过探索更精细的因果图模型和精准的干预,可以更好地理解推荐系统中的因果关系,为用户提供更加个性化和准确的推荐.

### 3.2.4 反事实推理

反事实推理的核心在于回答一种假设性语句,用来描述在不同的因果条件下,一个变量的值会是什么. 在第 3.2.3 节中,使用了等式  $X=x$  来代替结构方程中的  $X$  来模拟将变量  $X$  设置为值  $x$  的行为. 继续使用这种形式来描述反事实问题的话可以表述成:“假设实际情况下  $X=0$  时,观测值为  $Y_{X=0}(Y_0)$ ; 那么如果  $X=1$  时,观测值  $Y$  会如何变化?”下面介绍反事实的基本原理和估计反事实的方法.

对于任意一个结构模型  $M$ , 考虑任意两个变量  $X, Y$ . 令  $M_x$  表示用  $X=x$  代替  $X$  后得到的  $M$  的修改版,反事实  $Y_x(u)$  的形式化定义为

$$Y_x(u) = Y_{M_x}(u) \quad (6)$$

即模型  $M$  中的反事实  $Y_x(u)$  被定义为修改后的子模型  $M_x$  中  $Y$  的解. 一般来说,反事实遵循以下的一致性原则:

$$\text{IF } X=x \text{ THEN } Y_x=Y \quad (7)$$

如果  $X$  是二值的,那么一致性原则变为

$$Y = XY_1 + (1-X)Y_0 \quad (8)$$

反事实值的计算基本流程如下: (1) 溯因. 用证据  $E=e$  来确定  $U$  的值; (2) 作用. 修改模型  $M$ , 移除变量  $X$  出现在左边的方程,用  $X=x$  来替换他们,获得修正的模型  $M_x$ ; (3) 预测. 使用修正后的模型  $M_x$  和  $U$  值来计算  $Y$  的值,即为反事实的结果.

反事实同样可以使用因果图来形象化表示. 上述反事实原理说明,如果修改模型  $M$  得到子模型  $M_x$ , 那么修改的模型中的结果变量就是原始模型的

反事实  $Y_x$ . 因为模型修改要求移除所有指向  $X$  的箭头, 因此可以说, 与变量  $Y$  关联的变量替换为与  $Y_x$  关联, 这个替换只在修改后的模型中成立. 进一步可以得出后门的反事实解释, 即如果一个变量集合  $Z$  满足  $(X, Y)$  的后门条件, 那么对于所有的  $x$ , 在给定  $Z$  条件下, 反事实  $Y_x$  独立于  $X$ , 记为

$$P(Y_x | X, Z) = P(Y_x | Z) \quad (9)$$

这是由于给定了  $Z$ , 相当于在反事实的条件下切断了所有  $X$  与  $Y_x$  的联系. 以上是反事实推理的基本原理.

在推荐系统去偏的应用中, 反事实推理扮演着重要的角色. 一方面, 在推荐的数据层面, 反事实推理能够支持探索未观测的数据, 通过构造反事实空间数据, 不仅可以缓解推荐系统中的数据缺失问题并增强模型的鲁棒性, 还可以减轻特定任务中的虚假相关性, 从而减轻偏差<sup>[44-46]</sup>, 如位置偏差<sup>[47]</sup>. 另一方面, 在推荐的算法层面, 反事实推理可以指导推荐算法的调整, 使其在处理未观测或假设的用户行为时更加准确. 另外, 通过反事实推理还可以提高推荐系统的可解释性<sup>[48]</sup>. 综上所述, 反事实推理不仅有助于生成高质量的反事实数据, 而且有助于构建更加精准和公平的推荐算法, 从数据和算法两个层面提供去偏指导.

为了深入探讨如何有效地应对推荐系统中的偏差问题, 本文接下来将从因果推断的视角出发, 系统地分析推荐系统的去偏策略. 在本文第 2 节, 偏差被划分为数据阶段的偏差、算法阶段的偏差和评估阶段的偏差. 每一阶段的偏差都有着特定的挑战和解决思路, 基于上述因果技术, 下文第 4~6 节将分别介绍各阶段的去偏策略.

## 4 数据阶段的去偏方法

在数据阶段, 偏差主要来源于用户互动数据的收集过程. 本节从数据处理的角度来概述基于因果推断的推荐系统去偏方法, 目标是解决推荐系统数据缺失或不平衡的问题. 在这种场景下, 推荐系统的数据可以分为三类: 观测到的有偏数据、实验收集的无偏数据和未观测到的数据. 为了消除有偏数据里偏差的影响, 从数据阶段去偏的推荐主要有两种方式: 一类通过是构造反事实空间来估计未观测到的数据, 缓解数据缺失问题; 另一类是通过使用部分无偏数据辅助原始有偏数据的训练, 试图纠正原来的

数据不平衡问题. 接下来将分别介绍这两类方法.

### 4.1 基于反事实空间构造的去偏方法

推荐系统研究往往是建立在训练数据是完全随机缺失 (MCAR) 的假设上的, 然而在真实的世界中, 这种假设很难成立, 由于受到推荐系统的策略以及用户对项目的自主选择性等因素影响, 观测到的历史交互数据是一种非随机缺失 (MNAR) 数据. 这就导致了推荐系统的数据偏差. 缓解数据偏差的一种方法是进行反事实空间构造, 通过对已有的观测数据进行反事实推理, 生成一批反事实增强样本. 将反事实增强样本与原始样本结合在一起训练, 可以支持探索未观测的数据空间, 在尽量保证数据正确性的同时提高推荐的性能. 根据数据生成的类型, 基于反事实的空间构造一般可以分为学习评分以及增强序列两种形式.

首先举例说明推荐系统的评分缺失现象. 表 2 展示了 3 个用户对 4 个项目的评分数据, 其中  $\checkmark$  表示正向反馈,  $\times$  表示负向反馈,  $?$  表示未观测到的交互. 针对部分可观测数据导致的偏差问题, Dong 等人<sup>[49]</sup>详细讨论了基于反事实技术的策略, 以学习全样本空间模型. 例如, 文献<sup>[50]</sup>通过构建插补模型并学习反事实样本, 缓解选择偏差问题, 同时采用修正的双重鲁棒方法优化点击率预估问题; 文献<sup>[51]</sup>基于影响函数对样本进行调权, 使其分布与期望但未观测的数据分布接近. 这些研究表明, 通过构建反事实样本和修正有偏数据, 能有效帮助训练出无偏的推荐模型.

表 2 非随机缺失 (MNAR) 评分数据

	项目 1	项目 2	项目 3	项目 4
用户 a	$\checkmark$	$?$	$?$	$\times$
用户 b	$?$	$\checkmark$	$\times$	$?$
用户 c	$\times$	$?$	$\checkmark$	$?$

为了处理非随机缺失数据, Saito 等人<sup>[52]</sup>受无监督领域自适应中非对称三级训练框架<sup>[53]</sup>的启发, 提出了一种模型无关的元学习方法, 利用三个非对称的模型来互相学习和纠正. 其中一个主模型用于最终的评分预测, 两个辅助模型为未观测样本学习可靠的伪标签. 该算法在解决选择偏差时避免了倾向分数估计不准确和高方差的问题, 提出了一个不受倾向分数限制的新的理想损失函数上界, 提升了推荐质量. 唯一的不足是该算法需要训练三个模型, 增加了计算的复杂度和内存占用. Wang 等人<sup>[46]</sup>从信息论的角度出发, 基于信息瓶颈来进行反事实学

习,构建了一个反事实的变分信息瓶颈模型 CVIB,以解决非随机缺失数据的问题.该模型采用一个编码器作为评分预测模型,一个解码器作为反事实模型用来估计用户的反事实评分,使推荐模型和反事实模型达到一种相对平衡,有助于同时减少偏差和方差的影响.该工作首次将信息瓶颈理论应用于推荐系统的反事实学习,不仅能够有效减少偏差和方差,而且容易与其他推荐模型结合,具有很高的扩展性和鲁棒性. Yang 等人<sup>[54]</sup>将反事实增强应用到排序推荐任务中,设计了基于学习的方法生成更有信息量的反事实样本.该文分析了生成样本数量与模型预测误差之间的关系,提供了可靠的理论保证.

文献<sup>[55-57]</sup>利用反事实进行序列数据增强,缓解序列推荐中的数据稀疏和不平衡问题.为了从序列数据中学习更鲁棒的用户表征,文献<sup>[55]</sup>使用对比学习来对抗不必要或干扰性的概念.通过处理细粒度的项目级和更抽象的兴趣级概念,该文提供了更全面的用户理解.尽管在理论上进行了创新,但是此方法在实际应用中可能存在一定的难度,尤其是面对复杂或动态的用户行为数据时. Wang 等人<sup>[56]</sup>为了生成反事实的序列数据,设计了三种序列增强思路,包括一种简单的随机替换方法、一种面向数据的采样方法和一种面向模型的采样方法,为反事实数据增强提供了新的思路和方法.文献<sup>[55-56]</sup>侧重于增强用户的隐式反馈数据,而文献<sup>[57]</sup>用反事实推理来增强用户的显式反馈数据.他们提出了一种反事实协同推理方法 CCR 来生成增强的困难样本辅助训练,该方法生成的反事实样本能够帮助模型在面对不常见或未见过的用户行为时,仍然能够做出合理的推荐,增强了模型的泛化能力.以上三种序列推荐都是通过反事实生成信息更丰富的序列数据,缓解了原始数据的稀疏和不平衡性,同时 CCR<sup>[57]</sup>生成的显示反事实数据,具有更高的可解释性.

根据以上研究发现,基于反事实空间构造的去偏方法通过模拟和评估用户在不同条件下可能做出的决策,提供了对潜在用户行为的深入理解,同时有助于揭示和纠正数据中的潜在偏差.但是这种去偏策略需要复杂的模型来估计和生成反事实场景,计算成本可能较高.

#### 4.2 基于无偏数据的去偏方法

在推荐系统领域,无偏数据是指一种不受用户行为偏好影响的数据,通常是通过随机或者均匀的方式采集的,也称作均匀数据(uniform data).由于

这种数据未受选择偏差和曝光偏差的影响,因此能够真实反映用户的兴趣,而不仅仅是推荐系统的反馈效应.使用无偏数据可以提高推荐系统的多样性和公平性,避免推荐系统陷入过度个性化或者过度曝光的问题.

在基于无偏数据的去偏研究中,需要使用带有无偏数据的推荐系统数据集.这样的数据集尽可能减少了数据偏差,有利于研究者能够在更公平的条件下评估他们的推荐算法.常用的带有无偏数据的数据集有:Yahoo! R3 数据集<sup>[58]</sup>提供了随机推荐的音乐数据;Coat 数据集<sup>[59]</sup>提供了用户对随机展示商品的评分;KuaiRand 数据集<sup>[60]</sup>提供了用户对随机视频的真实反馈.因为收集无偏数据的成本很高,导致这些数据集比较稀缺.

研究者们提出了不同的策略来利用无偏数据增强推荐性能. Bonner 等人<sup>[61]</sup>最早提出使用部分随机数据进行去偏.他们提出了一种基于域适应策略的矩阵分解方法 CausE,并采用多任务学习方法来进行训练.该方法使用无偏数据和有偏数据联合学习两个模型,并使用正则项来调节两个模型的差异,其联合损失函数为

$$L_{\text{CausE}}^{\text{prod}} = \underbrace{L(\mathbf{U}\Theta_t, Y_t)}_{\text{无偏损失}} + \Omega(\Theta_t) + \underbrace{L(\mathbf{U}\Theta_c, Y_c)}_{\text{有偏损失}} + \underbrace{\Omega(\Theta_t - \Theta_c)}_{\text{正则项}} \quad (10)$$

其中,  $\Theta_t, \Theta_c$  分别是物品向量  $t, c$  的矩阵参数,  $\mathbf{U}$  是指用户向量的矩阵参数,  $L$  为损失函数,  $\Omega$  为正则项.该算法是对矩阵分解算法的扩展,首次提出利用一个小样本的无偏数据来提升模型的效果,其中随机数据是通过平衡物品曝光的概率产生的.该方法明显优于经典的矩阵分解算法.但是该方法比较简单,模型本身的通用性不高.

Liu 等人<sup>[62]</sup>进一步研究了随机数据的去偏性能.通过线上 A/B 测试和离线实验验证,即使只是简单地使用无偏数据,也能够缓解偏差问题,提高推荐的性能.验证结果如表 3 所示.

表 3 无偏数据的验证效果<sup>[62]</sup>

训练数据组成	离线 AUC 结果	线上 eCPM 结果
99%有偏数据	0.7571	0.0%
99%有偏+1%无偏数据	0.7689	2.98%(提升)

为了更有效地利用无偏数据,他们提出了一个通用的知识蒸馏框架 KDCRec 实现去偏推荐.该框架的核心是从无偏数据中挖掘有用的信息,来指导有偏数据模型的训练.如图 3 所示,为了从不同的角

度学习指导信息, 该文共设计了四种知识蒸馏方法. 这篇工作严谨地通过离线和在线两种方式验证了无偏数据的有效性, 并且提出的 KDCRec 是一种通用性很高的知识蒸馏框架, 能够从不同的角度利用无偏数据进行反事实推荐, 为相关研究者们提供了很好的研究思路.

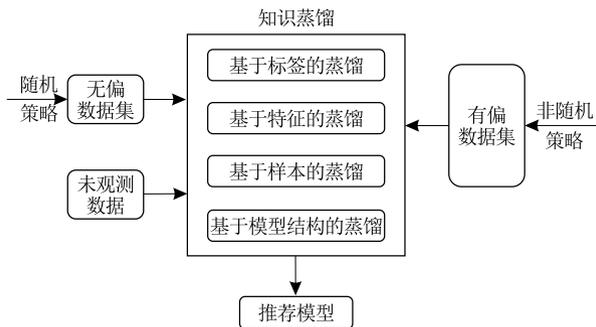


图 3 KDCRec 框架图<sup>[62]</sup>

上述两种方法, CausE 和 KDCRec, 为基于无偏数据去偏研究提供了充分的验证和研究基础. 为了增强去偏的性能, 研究者们提出了许多从不同的方面来利用无偏数据的方法. 例如, 文献<sup>[63]</sup>提出了一种新颖的目标函数, 使用无偏数据来动态调整偏见数据的倾向性权重. 该文提出的算法 LTD 可以最小化倾向性估计的方差, 从而更好地泛化推荐系统. 为了统一地消除推荐系统中的各种偏差, Chen 等人<sup>[64]</sup>创新性地从风险差异的角度出发, 将推荐系统中各种偏差的来源归结于预期经验风险和真实理想风险之间的差异, 通过减小风险差异推导出一个通用的去偏框架 AutoDebias. 作为一种自动去偏的方法, AutoDebias 不仅可以解决不同的偏差问题, 还可以应用到不同的推荐模型上, 具有很好的普适性. Li 等人<sup>[65]</sup>则首次提出使用少量无偏评分数据来平衡推荐数据中未观测的混杂因素. 他们的方法不仅使用无偏数据来校正模型参数, 还通过平衡权重进一步利用这些数据来减少预测偏误, 提高了无偏数据的使用效率.

以往的推荐方法一般只有在有偏或无偏一种测试集上进行评估. Ding 等人<sup>[66]</sup>实验发现许多去偏方法虽然在无偏测试集上效果明显, 但是在正常的有偏测试集上的性能却大大降低. 为了在两种数据上取得双赢, 该文试图构建一种对两类数据环境感知的推荐模型. 作者设计了一种基于标签的插值蒸馏框架 InterD, 从用户项目对级别对环境的评分分布进行细粒度插值. 该文提出的方法首次统一了有偏和无偏两种环境, 具有更高的泛化性和鲁棒性.

以上研究发现, 基于无偏数据的去偏方法直接利用已知无偏的数据作为训练基础, 不仅具有更加准确和高效的去偏效果, 而且相对于基于反事实空间构造的方法, 减少了对复杂模型的依赖, 相对容易实施. 但是这种去偏策略依赖高质量的无偏数据, 通常获取困难且成本较高.

### 4.3 其他方法

在推荐系统中, 还有一些较少被探讨但新颖的数据偏差, 如注意力偏差、时长偏差等, 影响了推荐的准确性和公平性. Zhang 等人<sup>[67]</sup>研究了用户注意力偏差在音乐流媒体推荐系统中的影响: 系统可能会过度重视那些被用户自动播放的歌曲, 导致学习到的用户偏好可能并不准确. 因此需要通过识别真实的用户行为和意图来校正注意力偏差. 他们设计了一种新颖的反事实学习方法, 通过引入神经决斗老虎机算法, 能够区分用户的真实反馈与因自动播放产生的假正反馈, 从而解决注意力偏差. 在视频推荐中, 时长偏差指长视频因为其时长较长, 往往能获得更长的观看时间, 误导推荐模型偏向推荐长视频, 从而忽视了用户的真实兴趣. Zheng 等人<sup>[68]</sup>对时长偏差进行了深入分析和验证, 提出了新的无偏评估指标“观看时长增益”和解决方案来减轻时长偏差, 更好地平衡推荐性能. 为了解决时长偏差, 文献<sup>[69]</sup>使用因果图分析了视频时长作为混杂因素对视频曝光率和观看时长预测的影响, 然后通过数据分割直接建模和消除时长偏差.

以上工作讨论了推荐系统中一些较为新颖的数据偏差, 不仅理论上具有创新性, 而且在流媒体等应用中具有实际效果, 对于提高视频推荐系统的用户参与度和满意度具有重要意义.

综上所述, 推荐系统数据阶段的去偏研究主要包含两种策略. 其中, 基于反事实空间构造的方法利用反事实技术构建并学习反事实样本, 但模型往往复杂性较高; 利用无偏数据的方法从事实的角度平衡数据偏差能够有效提高预测的准确性, 但是无偏数据量小且不容易获得. 从因果推断带来推荐结果的可解释性角度来看, 在数据阶段的去偏方法中, 可解释性主要体现在数据增强时的反事实推理和逻辑推理. 例如, CCR<sup>[57]</sup>通过对用户显示反馈进行最小化更改生成反事实数据, 这一过程提供了反事实解释, 同时模型引入神经逻辑模块为推荐提供了更深层的因果解释.

表 4 总结和分析了从数据阶段去偏的因果推荐方法.

表 4 数据阶段的因果推荐去偏方法

因果策略	去偏方法	推荐框架	应用场景	性能			文献来源	数据集
				鲁棒性	泛化性	解释性		
构造反事实空间	OCCTR <sup>[50]</sup>	LR/FM	选择偏差	✓	✓		CIKM 2019	Yahoo! R3/Coat
	IF4URec <sup>[51]</sup>	FFM	选择偏差	✓	✓		SIGIR 2020	Yahoo! R3/CPC
	AT <sup>[52]</sup>	MF-based	选择偏差	✓			SIGIR 2020	Yahoo! R3/Coat
	CVIB <sup>[46]</sup>	MF/NCF	选择偏差	✓	✓		NeurIPS 2020	Yahoo! R3/Coat
	CPR <sup>[54]</sup>	BPR/GMF/ MLP/NeuMF/ LightGCN	混合偏差	✓	✓		CIKM 2021	MIND
	Causerec <sup>[55]</sup>	MLP	位置偏差	✓	✓		SIGIR 2021	Books/Yelp/Gowalla
	CASR <sup>[56]</sup>	RNN/CNN/ Transformer	混合偏差	✓	✓		SIGIR 2021	MovieLens/Diginetica
使用无偏数据	CCR <sup>[57]</sup>	RNN/CNN/ Transformer	混合偏差	✓	✓	✓	WSDM 2023	ML100K/Amazon
	CausE <sup>[61]</sup>	MF	混合偏差	✓	✓		RecSys 2018	MovieLens10M/Netflix
	KDCRec <sup>[62]</sup>	MF/AE	混合偏差	✓	✓	✓	SIGIR 2020	Yahoo! R3/Product
	LTD <sup>[63]</sup>	MF/NF	选择偏差	✓	✓		WSDM 2021	Music/Coat
	AutoDebias <sup>[64]</sup>	MF/FM/NCF	混合偏差	✓	✓		SIGIR 2021	Yahoo! R3/Coat
	Bal-model <sup>[65]</sup>	MF/NCF	混合偏差	✓	✓	✓	WWW 2023	Music/Coat
其他方法	InterD <sup>[66]</sup>	MF	混合偏差	✓	✓		SIGIR 2022	Yahoo! R3/Coat/Product
	NBD <sup>[67]</sup>	GRU FM/WDL/	注意力偏差	✓			KDD 2022	Last.fm
	DVR <sup>[68]</sup>	DeepFM/NFM/ AutoInt/AFN	时长偏差	✓	✓		MM 2022	Wechat/Kuaishou
	D2Q <sup>[69]</sup>	MLP	时长偏差		✓		KDD 2022	Kuaishou

## 5 算法阶段的去偏方法

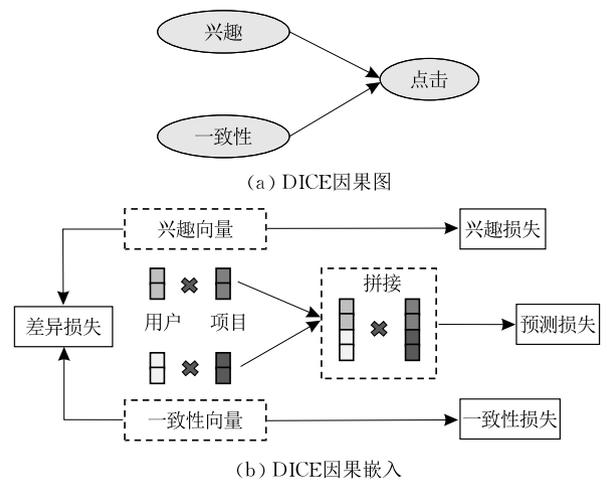
在算法阶段,偏差主要来源于模型学习过程中的不足.本节将介绍如何运用因果推断框架,通过改善算法的学习机制来减少偏差,从而达到更准确的推荐.算法阶段的去偏方法主要包含三类:基于因果表示学习的方法、基于因果干预的方法和基于反事实推理的方法.

### 5.1 基于因果表示学习的去偏方法

因果表示学习利用机器学习方法从观测数据中进行因果建模和表示学习.在推荐系统中,基于因果表示学习的方法一般是先用因果图对推荐系统进行因果建模,然后从向量的角度直接干预混杂因素,从而学习无偏的用户或项目向量进行推荐.基于因果表示学习的方法不仅能准确地建模用户的喜好,同时具有更高的鲁棒性和泛化性.目前基于因果表示学习的去偏推荐主要有学习解耦的因果表示和学习在不同数据环境中不变的因果表示两种思路.

在推荐系统中,解耦表示学习通过对用户和项目进行更细粒度的分析,旨在发现和解开隐藏在观测数据中潜在的影响因素.基于因果分析的解耦表示学习能够分离与用户偏好相关的因素和其他混杂因素,为去偏研究提供了一个很好的方案.文献[70]是最早提出从用户隐反馈中解耦因果表示的方法.他们通过因果性分析,将用户对物品的选择归因于两个

独立的因素:用户的兴趣和用户一致性.根据因果图的对撞结构,提出了一种新的解耦框架 DICE.如图 4 所示,该框架利用单独的嵌入来表示用户兴趣和项目受欢迎程度,并通过针对特定原因的训练样本和直接解耦监督来训练模型.Zhao 等人<sup>[71]</sup>提出的 DCCL 算法沿用了文献[70]中的因果假设,但是在解耦技术的应用和实现方法上有所不同.DICE 侧重于通过直接监督和因果推断明确分离不同因素的影响,而 DCCL 则通过对比学习来强化模型的数据表示能力,尤其在处理数据稀疏性方面表现出独到的优势.

图 4 DICE 框架图<sup>[70]</sup>

文献[72]和[73]是两个利用表示学习解决流行度偏差的工作.文献[72]分析了流行度偏差的两个

主要原因:一是推荐模型本身的训练范式为了获得更小的损失更容易推荐热门物品,二是长尾物品的稀疏性使它们的表示学习更困难.他们设计了一个基于域适应的解耦网络 CD2AN 同时训练有偏和无偏模型. CD2AN 和 DICE<sup>[70]</sup>、DCCL<sup>[71]</sup>不同的是,后者将用户兴趣和一致性从用户嵌入中分离出来,而前者将流行度信息与项目属性嵌入分离.文献[73]设计了两种训练策略来中和流行度偏差,学习与流行度无关的项目嵌入.文章直接在向量嵌入上消除流行度偏差的策略可灵活插入不同的推荐模型,在保持整体推荐准确性的同时减少流行度偏差,具有很高的鲁棒性和适用性.

在解耦表示学习中,变分自编码器(VAE)也是一种常用的技术.文献[74-75]均使用 VAE 作为基础架构,利用其生成模型的能力来学习用户行为数据的解耦表示,从而提高推荐的准确性和可解释性.其中,CAD-VAE<sup>[74]</sup>专注于通过结构因果模型来解耦潜在因素之间的因果关系,而 CDR<sup>[75]</sup>使用对比学习来强化潜在空间中因果关系的表示,使模型能够更好地区分不同因素对用户选择的影响,提高了模型在处理稀疏数据时的鲁棒性.

经典的推荐系统都是建立在独立同分布(i.i.d)假设的基础上,即训练数据和测试数据是独立同分布的.然而在真实场景中,该假设很难满足,导致推荐算法在分布变化时性能下降,也就是分布外(Out-of-Distribution, OOD)泛化问题.不变的因果表示学习不受与环境因素影响的偏好表示. Wang 等人<sup>[76]</sup>认为推荐系统里的 OOD 现象源于用户特征发生变化导致数据分布偏移,提出了一个新颖的推荐框架来处理用户兴趣特征转移.在文献[77]中,假设用户行为是由用户兴趣决定的不变偏好和受环境混杂因素共同决定的,他们提出了一种不变表示学习方法 InvPref 将用户不变偏好与潜在偏差区分开来.不变的因果表示学习和解耦的因果表示学习一样具有很好的鲁棒性,同时泛化性更高.

以上研究发现,基于因果表示学习的去偏方法通过挖掘和利用潜在的因果关系,提取潜在的因果特征,不仅增强了模型对用户行为的深入理解,而且增强了模型的鲁棒性和泛化性.但是因果表示学习通常涉及复杂的数据处理和模型训练,面临计算成本高和对数据的依赖性强等挑战.

## 5.2 基于因果干预的去偏方法

基于因果干预的推荐方法使用因果图构建变量之间的因果关系,通过后门调整或前门调整技术来消除混杂因素,从而消除偏差的影响.和 5.1 节不

同,基于因果干预的方法侧重于使用后门调整或前门调整技术从更宏观的因果关系角度消除混杂因素的影响.这些方法不仅关注于单一的向量调整,而且从整个系统的因果结构出发,通过识别并调整影响用户行为的根本因果路径,以达到去偏的目的.

Zhang 等人<sup>[78]</sup>认为推荐系统建模时的流行度偏差并非全是有利的.为了消除流行度偏差有害影响的同时利用其好的影响,作者构建了一个基于因果干预来平衡流行度偏差的模型 PDA.首先,如图 5 所示,使用因果图描述了推荐数据的生成过程:边  $\{U, I, Z\} \rightarrow C$  表示交互标签  $C$  由用户  $U$ 、项目  $I$  和项目流行度  $Z$  三种因素共同决定,边  $Z \rightarrow I$  表示项目流行度影响了项目的曝光.也就是说在这一过程中,一方面,流行度能够直接影响项目的曝光频率;另一方面,用户因从众心理会增大与流行度高的物品的交互概率.因此,流行度可以被视为一种同时作用于曝光项目  $I$  和交互标签  $C$  的混淆因素,它导致从  $Z$  通向  $C$  存在两条路径: $Z \rightarrow C$  和  $Z \rightarrow I \rightarrow C$ .在训练时,作者利用  $do$  算子进行概率推导并简化模型,消除流行度偏差的影响;在预测时,利用预测的物品流行度,影响模型的排序结果.在这篇文章中,作者构建的因果图模型非常的清晰,并进行了有效的推导和模型简化,是一种非常经典的基于因果干预的去偏方法.文献[79]同样使用因果干预处理流行度偏差.他们研究了用户多行为推荐场景下的交互生成因果关系,利用后门调整阻断了混杂因素造成的后门路径,提高了推荐结果的公平性.

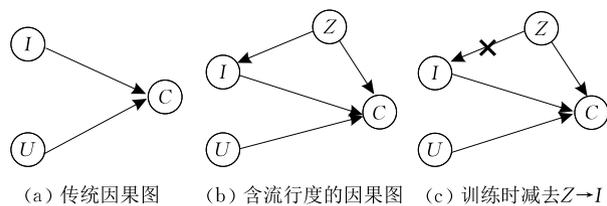


图 5 PDA 因果图建模过程<sup>[78]</sup>

文献[80]使用因果干预的手段处理反馈循环偏差.该文首先从因果的角度分析了偏差放大的原因,并使用因果图刻画了推荐过程中的虚假相关关系,如图 6 所示.其中,  $U$  为用户表示,  $I$  为项目表示,  $D$  表示用户在项目群体上的分布,  $M$  表示群体级别的用户表示,  $Y$  为预测评分.接下来提出了一个去混杂效应推荐系统 DecRS,它模拟了用户表示对预测分数的因果效应.消除混杂因素影响的关键在于后门调整,即  $P(Y|do(U=u), I=i)$ ,其中,  $do(U=u)$  在因果图中表现为剪掉的边  $D \rightarrow U$ ,从而阻断了  $D$  对  $U$  的影响,后门调整的推导公式为

$$\begin{aligned}
& P(Y|do(U=u), I=i) \\
&= \sum_{d \in \mathcal{D}} P(d|do(U=u)) P(Y|do(U=u), i, \\
&\quad M(d, do(U=u))) \\
&= \sum_{d \in \mathcal{D}} P(d) P(Y|do(U=u), i, M(d, do(U=u))) \\
&= \sum_{d \in \mathcal{D}} P(d) P(Y|u, i, M(d, u)) \quad (11)
\end{aligned}$$

DecRS 是少数解决推荐系统偏差放大问题的研究工作之一,该工作描述的因果关系推理和基于后门调整的计算也都非常完善。

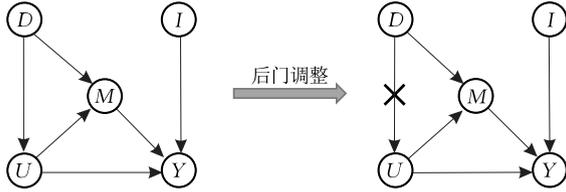


图 6 DecRS 因果图建模<sup>[80]</sup>

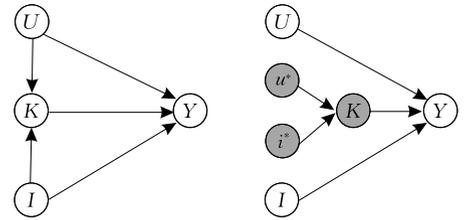
文献[81]使用因果推断解决情感偏差问题,他们将评论文本呈现的情感极性作为用户/项目表示和观测评分之间的混杂因素,通过干预手段消除情绪偏差的负面影响。这是一种基于评论信息的去偏推荐方法,通过因果干预的手段消除了推荐系统的情感偏见,同时平衡了部分情感倾向实现个性化推荐。Wang 等人<sup>[82]</sup>提出了一种因果启发式的干预方法 CausalInt 来处理多推荐场景之间的互相干扰问题,通过因果干预来建立场景感知信息和用户行为之间的因果关系,并将其与场景之间的不变表示联系在一起。为了同时解决推荐系统中遇到的各种偏差, Xu 等人<sup>[83]</sup>提出了一个去混杂的因果协同过滤模型 DCCF。首先将用户行为与未观察到的混杂因素构成因果图,然后设计一个前门调整模型,结合机器学习,以消除未观察到的混杂因素的影响。考虑到在真实系统中项目数量规模很大,因此 DCCF 设计了一个基于采样的方法结合进暴露模型,使得在项目空间很大的情况下让前门调整可计算。该模型既能处理全局混杂因素,也能处理个性化混杂因素。相似地,文献[84]提出了一个去混杂因果推荐方法 DCR 对虚假关联进行后门调整。为了方便  $do$  运算评估,他们设计了一个专家模块对混杂特征的值进行建模。

以上工作设计了不同的因果图来分析不同偏差的生成过程,然后再基于后门调整消除混杂因素。大多数基于因果干预的去偏方法难以验证其假设的因果关系是否正确,难点在于进行后门调整的概率推导时,样本空间的分布很难计算,需要设计特定的策略,如文献[80,83-84]。

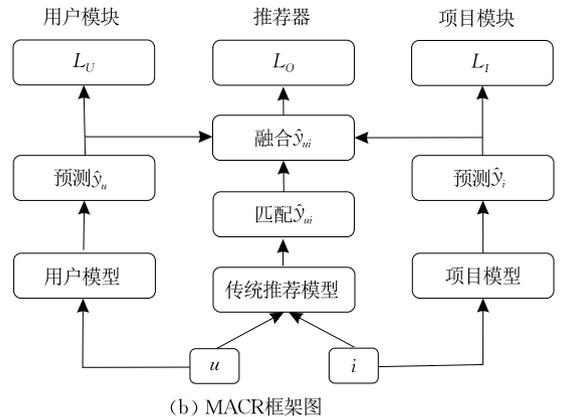
### 5.3 基于反事实推理的去偏方法

基于反事实推理的推荐系统通过对用户行为数据进行因果推断,来模拟用户在不同情况下的行为,从而提高推荐系统的效果。例如,通过反事实模拟用户在没有看到某个物品时会选择哪个物品,可以更好地理解用户的偏好。这种方法可以有效地解决数据稀疏性问题,并提高推荐系统的效果。

和文献[70-72,78]类似,文献[85]从因果推理的角度解决流行度偏差问题。作者假设有三种主要因素影响了用户-物品的交互概率:用户-物品的匹配程度、用户的从众程度以及物品本身的流行程度,并依此构建了因果图,如图 7(a)所示。为了消除流行度的影响,文中设计了一个反事实推理的框架 MACR,如图 7(b)所示,来评估“用户-物品的匹配程度”对“评分”的因果效应,将预测评分减去因果图中偏差造成的影响即得无偏的预测分数,从而实现反事实推理。该文的亮点在于使用因果图构建因果关系,建立了物品对评分的直接影响,进而使用反事实推理消除了物品的流行度属性对评分的影响。



(a) 事实与反事实因果图



(b) MACR 框架图

图 7 MACR 推理与实现过程<sup>[85]</sup>

因果推断最基础的概念之一是相关关系不等于因果关系,相关性可以直接地从观测数据中发现而因果性则更难识别。虚假相关性指的是两个相关的变量之间实际上并不存在直接的因果关系。在基于知识感知的推荐系统中,虚假相关性常由那些与用户交互频繁出现但非因果的项目特征引起,如购买或点击行为。由于集中于观测到的共现模式,这些方法

很容易被虚假联系误导,并且很难去识别真正的因果联系。Mu 等人<sup>[86]</sup>提出构建反事实交互,在知识感知的推荐中识别和减弱潜在虚假相关性。为了缓解潜在虚假相关性,他们通过强化学习生成了含有虚假的正向和负向反馈的数据。这是第一个从反事实推理的角度来消除虚假关联的工作,具有一定的借鉴意义。

Wang 等人<sup>[87]</sup>重点关注了推荐系统中常见的虚假点击现象,即用户可能因为广告的噱头而点击了广告,但发现实际内容并非自己所需要的。在这样的数据上进行模型训练,很有可能导致马太效应,使得模型倾向于一些骗取点击的流量。这种现象被称作“点击诱饵偏差”,属于数据偏差的一种。作者利用因果图进行因果推断,估计在反事实的情况下用户的点击和消费情况来识别和缓解骗取点击的流量。文章的整体思路是通过从实际点击数据中剔除掉曝光特征的直接效应,来估计物品整体特征带来的实际效果。

He 等人<sup>[88]</sup>利用观测数据获取用户和项目混杂因素的替代品,通过反事实推理来消除混杂因素的因果效应,从而实现一种通用的去偏差框架 CIDR。文献<sup>[89]</sup>旨在解决部分混杂变量不可观测,导致 IPS 方法在推荐系统中应用时不满足可识别性原理的问题。作者提出了一种最小最大经验风险框架,利用两个推荐模型来模拟松弛后的对偶问题,进而解决曝光偏差对预估推荐结果的影响。文章证明了所提出的对抗性学习的泛化界限,并为极大极小值优化提供分析,其思路可以适用于多种场景。文献<sup>[90]</sup>提出利用反事实的方法来建模用户的评论行为。作者通过干预用户的偏好来生成反事实样本,使用观测样本和反事实样本共同训练推荐模型,提升模型性能。和模型 DCCF<sup>[83]</sup>相似,都是先利用反事实本来增强模型,生成方法主要是基于学习的方法,目标是生成难样本来最大限度地提升模型的性能。

为了解决大规模 Top-K 推荐系统中数据偏差导致的策略学习问题,如重要性权重爆炸、观察不足和训练效率等,文献<sup>[91]</sup>提出了一个实用的基于反事实学习框架。该框架设计了倾向性矫正、自适应策略优化和高效训练模块来应对上述挑战。这篇文章对基于反事实的策略学习方法做了非常充足的研究与讨论,所提出的方法实用性和效率很高,不足之处在于只在 MovieLens 数据集上做了验证,缺乏在不同数据集上的泛化性验证。Agarwal 等人<sup>[44]</sup>为基于反事实的排序学习提供了一个通用且理论上严谨的框架,推导出了一个松弛后的可微分的基于倾向加权的排名指标,使得目标函数可以采用梯度下降的方法进行求解,以此来学习模型参数。基于该框架的

模型可以用来优化可加排序值,以此进行模型的无偏学习,从而解决数据偏差的问题。

在推荐系统去偏领域,基于因果推断的不公平性偏差问题往往通过反事实推理的方法处理。Huang 等人<sup>[92]</sup>研究了如何通过将因果推理纳入 bandits 中,实现用户在线推荐中的反事实公平。作者采用软干预对 arm 选择策略进行建模,并使用从因果图中识别的  $d$ -分离开发了一种公平的 UCB 算法,该算法通过选择满足反事实公平约束的 arm 来促进公平。文献<sup>[93]</sup>研究了如何通过反事实推理来解释哪些特征导致推荐中的项目暴露不公平的问题。因果推断的关键目标是调查数据和模型背后的关系,包括敏感变量和决策之间的因果效应,以及敏感变量和非敏感变量之间的相关性。通过切断敏感特征与输出结果之间的因果关系来实现反事实公平。Lin 等人<sup>[94]</sup>研究了群体推荐问题。作者通过考虑每个推荐项目与用户的相关性来定义向小组建议的个人效用,并将组成员的效用设施之间的不平衡视为公平性指标。作者最大限度地提高每个小组成员的满意度,同时尽量减少他们之间的不公平。Li 等人<sup>[95]</sup>考虑建议的反事实公平性,这要求每个用户的建议结果在事实和反事实世界中是相同的。反事实世界被定义为用户的敏感特征被改变,而所有其他不因果关系于敏感特征的特征保持不变。基于反事实的公平除了上述中常用的基于因果推断中的干预和推理以及直接移除敏感属性等手段,还包括变分编码器、对抗学习、数据增强等。

根据以上研究工作发现,基于反事实推理的方法在去偏研究内容广泛,在去数据偏差、混杂因素、虚假相关性以及不公平性偏差等问题上取得了非常多的成果。这些方法通过构建反事实场景,可以在理论上探索不同推荐方案的潜在影响,增强推荐的可靠性和科学性。但是有效的反事实推理依赖于准确的模型和假设,错误的模型设定可能导致误导性的结论,且具有一定的实施难度。

#### 5.4 其他方法

对于推荐系统中存在的未观测混淆因素,有相关工作提出了新的因果解决思路。例如,文献<sup>[96]</sup>分析了推荐系统中由于未测量变量导致的混淆偏差问题,引入了代理变量和近因因果推断技术,提高了推荐模型在存在未测量混淆因素时的准确性和鲁棒性。文献<sup>[97]</sup>采用了因果图模型来形式化推荐系统中的因果关系,并利用敏感性分析来量化未测量混淆因素对推荐结果的影响。文献<sup>[98]</sup>探讨了如何将深度学习与因果推理相结合,利用因果推理来理解

和解决推荐过程中的各种偏差,同时利用深度学习模型的强大表示能力来捕捉复杂的用户行为模式和偏好.以上工作从不同的角度为基于因果的推荐系统去偏研究提供了新的视角和解决方案.

除此之外,基于因果的去偏研究也出现了一些新的推荐应用场景.新奇项推荐旨在向用户推荐那些偏离他们以往行为和预期的物品,目的是通过引入新颖性来打破用户的惯性和过滤气泡,从而增加用户的满意度和参与度<sup>[99-100]</sup>.文献[100]因果推断应用到新奇项推荐系统的开发中,使用潜在因果框架解决如何确定推荐新奇项的最佳时机以及推荐的新颖度的挑战.

综上所述,推荐系统算法阶段的去偏研究主要包含三种因果策略.其中,基于因果表示的去偏方法从向量的层面干预混杂因素,学习用户的因果表示,具有较高的鲁棒性和泛化性;基于因果干预的去偏方法利用因果图对推荐的生成过程进行审查,假设性和理论性强同时需要复杂的推导和计算;基于反

事实推理的去偏方法通过设计反事实推理模型来回答假定的反事实问题,消除变量之间的虚假相关性.可以发现,算法阶段的三种去偏策略进一步增强了推荐系统的可解释性,它们往往借助于因果图来分析推荐决策背后的因果机制.例如,图 4(a)展示了 DICE<sup>[70]</sup>将用户意图解耦为用户兴趣和流行趋势,从数据的角度提升了推荐的可解释性;DecRS 分析了推荐系统偏差放大的原因,通过图 6 刻画了推荐过程中的虚假相关关系,利用后门调整消除了混杂因素的影响<sup>[80]</sup>;如图 7(a)所示,MACR 使用因果图建立了用户-物品匹配度、用户从众性、物品流行度与评分之间的因果关系,通过反事实推理消除了流行度偏差的影响<sup>[85]</sup>.这些基于因果图模型的推荐方法明确了用户行为与项目之间的因果路径,从而避免混淆变量引入的偏差,不仅提高了推荐结果的准确性,还提升了推荐系统的可解释性.

表 5 总结和分析了从算法阶段去偏的因果推荐方法.

表 5 算法阶段的因果推荐去偏方法

因果策略	去偏方法	推荐框架	应用场景	性能			文献来源	数据集
				鲁棒性	泛化性	解释性		
基于因果表示学习去偏	DICE <sup>[70]</sup>	MF/GCN	一致性偏差	✓	✓	✓	SIGIR 2022	MovieLens-1M/Netflix
	DCCI <sup>[71]</sup>	MF/LightGCN	流行度偏差	✓	✓	✓	WWW 2023	Yelp/Short-video
	CD <sup>2</sup> AN <sup>[72]</sup>	MLP	流行度偏差	✓	✓	✓	SIGIR 2022	MovieLens-1M/Taobao
	PID <sup>[73]</sup>	BPR/LightGCN	流行度偏差	✓	✓	✓	SIGIR 2022	ML-100K/Amazon DM
	CaD-VAE <sup>[74]</sup>	VAE	混合偏差	✓	✓	✓	arXiv 2023	ML-1M/Netflix
	CDR <sup>[75]</sup>	VAE	混合偏差	✓	✓	✓	TOIS 2023	Yelp/Amazon Book/Amazon Electronic
	COR <sup>[76]</sup>	VAE	混合偏差	✓	✓	✓	WWW 2022	Meituan/Yelp
基于因果干预去偏	InvPref <sup>[77]</sup>	MF	混合偏差	✓	✓	✓	SIGKDD 2022	Yahoo! R3/Coat
	PDA <sup>[78]</sup>	LFM	流行度偏差	✓	✓	✓	SIGIR 2021	Kwai/Douban/Tencent
	MBD <sup>[79]</sup>	MF	流行度偏差	✓	✓	✓	TCSS 2024	Kwai/Tmall
	DecRS <sup>[80]</sup>	FM/NFM	反馈循环偏差	✓	✓	✓	SIGKDD 2021	ML-1M/Amazon Book
	CISD <sup>[81]</sup>	LFM/MLP/FM	情感偏差	✓	✓	✓	CIKM 2022	Amazon Gourme/Kindle/Electronics/Video Games/Yelp
	CausalInt <sup>[82]</sup>	MLP	场景偏差	✓	✓	✓	SIGKDD 2022	Ali-CCP/Industrial
	DCCP <sup>[83]</sup>	MLP	曝光偏差	✓	✓	✓	TORS 2023	Electronics/CDs and Vinyl/Yelp
基于反事实推理去偏	DCR <sup>[84]</sup>	NFM	混合偏差	✓	✓	✓	TOIS 2023	Kwai/Wechat
	MACR <sup>[85]</sup>	MF/LightGCN	流行度偏差	✓	✓	✓	SIGKDD 2021	Adressa/Globo/ML10M/Yelp 2018/Gowalla
	CGKR <sup>[86]</sup>	GNN	混合偏差	✓	✓	✓	SIGIR 2022	Yelp/Music/Movie
	CR <sup>[87]</sup>	MMGCN	点击诱饵偏差	✓	✓	✓	SIGIR 2021	Tiktok/Adressa/
	CIDR <sup>[88]</sup>	MMGCN	混合偏差	✓	✓	✓	ECML-PKDD 2022	Tiktok/Adressa/Coat
	ACL <sup>[89]</sup>	GMF/ATTN	曝光偏差	✓	✓	✓	NeurIPS 2020	ML-1M/LastFM/Goodreads
	CP <sup>2</sup> <sup>[90]</sup>	ATTN	混合偏差	✓	✓	✓	CIKM 2021	Digital Music/Tools&.Home/Home&.Kitchen/Office Products/Yelp
	RIIPS <sup>[91]</sup>	Policy learning	选择偏差	✓	✓	✓	SIGKDD 2022	ML-1M/ML-10M
	L2R <sup>[44]</sup>	SVM	位置偏差	✓	✓	✓	SIGIR 2019	Yahoo/LETOR4.0
	F-UCB <sup>[92]</sup>	UCB	不公平性偏差	✓	✓	✓	AAAI 2022	Adult-Video
其它方法	CEFs <sup>[93]</sup>	MLP	不公平性偏差	✓	✓	✓	SIGIR 2022	Yelp/Electronics/CDs&.Vinyl
	Greedy-Var <sup>[94]</sup>	—	不公平性偏差	✓	✓	✓	RecSys 2017	MovieLens/MoviePilot
	iDCF <sup>[96]</sup>	MF/VAE	混合偏差	✓	✓	✓	KDD 2023	Yahoo! R3/Coat/KuaiRand
	BRD <sup>[97]</sup>	IPS/DR/AutoDebias	混合偏差	✓	✓	✓	KDD 2022	Yahoo! R3/Coat/Product
Deep-Deconf <sup>[98]</sup>	DNN	混合偏差	✓	✓	✓	arXiv 2022	Simulated/ML-causal/VG-causal	

## 6 评估阶段的去偏方法

评估阶段的偏差影响了对推荐系统性能的正确理解. 本节将讨论如何利用因果推断方法改进评估指标和流程, 以更真实地反映系统的效果. 在评估阶段的去偏中, 同样含有选择偏差等数据偏差场景, 但是与数据阶段的去偏方法有所不同: 数据阶段的去偏侧重于从数据处理的层面解决原有的数据缺失问题, 而评估阶段的方法侧重于在已知训练数据是缺失的前提下, 修正和优化评估函数, 以获得相对无偏的估计.

对于推荐系统来说, 能够准确地预测用户对物品的反馈, 这在排序阶段至关重要. 因此, 需要能够准确地衡量每个用户对每个物品的评分是否足够准确. 由于用户只会对喜爱的物品进行评分, 很少对不喜欢的物品进行评分, 导致评分矩阵是稀疏的, 并且不具有代表性, 常用的经验风险最小化估计方法并不是对所有用户对物品的评分的无偏估计, 这产生了评估偏差.

### 6.1 基于逆倾向性评分的去偏方法

Schnabel 等人<sup>[59]</sup>首次将逆倾向评分应用在无偏模型评估. 在一个电影推荐系统中, 用户只会对他们喜欢的电影观看和评分, 并且很少对不喜欢的电影评分. 以想要优化的效果为条件进行观察会导致数据是非随机缺失的<sup>[101-102]</sup>. 作者认为在推荐系统中将物品曝光给用户相当于在医学诊疗中对病人采用某种疗法. 对于推荐系统来说, 想要根据真实的评分  $\mathbf{Y}$  来评估预测出的结果  $\hat{\mathbf{Y}}$  有多好, 一般采用经验风险最小化的评估方法:

$$R(\hat{\mathbf{Y}}) = \sum_{u=1}^U \sum_{i=1}^I \delta_{u,i}(\mathbf{Y}, \hat{\mathbf{Y}}) \quad (12)$$

其中  $u$  和  $i$  分别表示用户和项目,  $\mathbf{Y}$  和  $\hat{\mathbf{Y}}$  分别表示真实标签和预测标签. 这是一种理想情况下的评测指标, 是建立在所有用户-物品对的评分全部可观测的假设之上, 但是如表 2 所示, 只能观察到评分矩阵  $\mathbf{Y}$  中的一部分, 传统方式采用已观测评分的平均值来评估  $R(\hat{\mathbf{Y}})$ :

$$\hat{R}_{\text{naive}}(\hat{\mathbf{Y}}) = \frac{U \cdot I}{|\mathbf{O}_{u,i}=1|} \sum_{(u,i); \mathbf{O}_{u,i}=1} \delta_{u,i}(\mathbf{Y}, \hat{\mathbf{Y}}) \quad (13)$$

其中,  $\mathbf{O}$  表示观测数据指示矩阵, 当真实评分  $r_{u,i}$  被观测到时,  $\mathbf{O}_{u,i}=1$ ; 当  $r_{u,i}$  缺失时,  $\mathbf{O}_{u,i}=0$ . 由于选择偏差的存在, 导致已观测评分并不能反映真实的评

分分布, 因此这种估计方法是有偏的, 为了解决选择偏差造成的评估偏差, Schnabel 等人<sup>[59]</sup>基于倾向得分匹配 (Propensity Score Matching) 的方法对观测到的结果进行加权调整, 得到了无偏的估计方法 IPS:

$$\hat{R}_{\text{IPS}}(\hat{\mathbf{Y}}|P) = \frac{1}{U \cdot I} \sum_{(u,i); \mathbf{O}_{u,i}=1} \frac{\delta_{u,i}(\mathbf{Y}, \hat{\mathbf{Y}})}{P_{u,i}} \quad (14)$$

其中  $P_{u,i}$  是定义在一个已观测评分上的边际概率, 可以通过朴素贝叶斯、逻辑回归等模型进行预测.

SNIPS (Self-Normalized Inverse Propensity Scoring)<sup>[59]</sup>是在 IPS 的基础上改进的一种方法, 目的是解决 IPS 方法方差过大的问题, 具体而言, SNIPS 在 IPS 的权重计算公式中引入了一个正则化项, 用于平衡观测数据的权重:

$$\hat{R}_{\text{SNIPS}}(\hat{\mathbf{Y}}|P) = \frac{\sum_{(u,i); \mathbf{O}_{u,i}=1} \delta_{u,i}(\mathbf{Y}, \hat{\mathbf{Y}}) / P_{u,i}}{\sum_{(u,i); \mathbf{O}_{u,i}=1} 1 / P_{u,i}} \quad (15)$$

相对于 IPS, SNIPS 估计器的方差更小, 代价是引入了小部分偏差.

Joachims 等人在他们的工作<sup>[103]</sup>中提出一种新的偏差——位置偏差, 并利用倾向性权重给出了理想损失函数的无偏估计. 隐式反馈数据虽然易于收集, 但是它自带位置偏差, 也就是说排序越靠前的物品, 用户越容易浏览到并点击. 作者使用倾向性评分对损失函数加权, 受到位置偏差影响越大的物品 (排序靠后仍被点击的物品) 权重越高. 对于一个排序模型来说, 理想的损失函数是

$$\theta^* = \arg \min \mathcal{L}(\theta) = \arg \min_{\theta} \int_{\mathcal{Q}} l(f_{\theta}, r_q) dP(q) \quad (16)$$

也就是打分函数  $f$  与真实的相关性  $r$  的误差, 但是目前的排序模型损失函数是

$$\begin{aligned} \mathcal{L}'(\theta) &= \int_{\mathcal{Q}} \int_{\pi_q} \mathbb{E}_{\mathbf{c}_{\pi_q}} [l'(f_{\theta}, \mathbf{c}_{\pi_q})] dP(q, \pi_q) \\ &= \int_{\mathcal{Q}} \mathbb{E}_{\mathbf{c}_{\pi_q}} [l'(f_{\theta}, \mathbf{c})] dP(q, \pi_q) \end{aligned} \quad (17)$$

其计算的是打分函数  $f$  与点击  $\mathbf{c}$  的误差, 但是点击数据包含着位置偏差的影响, 作者采用倾向性评分, 对单个文档的损失加权:

$$l'(f_{\theta}, \mathbf{c}) = l_{\text{IPW}}(f_{\theta}, \mathbf{c}) = \sum_{d, c_d=1} \frac{\Delta(d, c_d | f_{\theta})}{P(\mathbf{O}_d=1)} \quad (18)$$

用户往往是从前到后浏览文档的搜索结果, 如果一个排名靠后的文档被点击, 说明这个文档对用户很重要, 因此根据受偏向影响的程度加权, 靠后的被点击文档权重重大, 靠前位置的被点击文档权重小, 被点击的文档受到影响越大, 权重给的越多, 模型越

关注这个样本。

Ai 等人<sup>[104]</sup>发现从点击数据中无偏倾向分数估计和无偏排序是对偶问题,并以此提出了 DAL 以同时学习无偏的排序模型和偏好分数模型。Saito 等人<sup>[105]</sup>使用倾向得分截断方法改进 IPS,降低无偏损失函数的方差。当倾向分数定义为曝光概率时,长尾物品的方差会很大。因此,作者采用了倾向得分截断方法平衡偏差和方差,并证明即使截断了倾向得分,仍能保证消除了偏差,并降低方差。文献[106]考虑到了即使推荐系统没有将物品曝光给用户,用户也会点击。因此,以物品评分作为排序的依据是不合适的。作者希望推荐由推荐系统导致点击的物品,这可以表示为估计曝光这一行为的因果效应问题。作者基于逆倾向评分技术构建了无偏的排名估计器指标,对估计量最小化经验风险的上界,减少有限训练样本下的方差。Lee 等人<sup>[107]</sup>为了消除曝光偏差,提出了 BISER,由自逆倾向权重和双向无偏学习两个部分组成,自逆倾向权重逐步消除物品曝光偏差,双向无偏学习用于弥补基于用户的以及基于项目的两个自动编码器之间的差距,消除自逆倾向权重的高方差。Li 等人<sup>[108]</sup>发现现有的 IPS 类方法缺乏对倾向分数统一且明确的标准,提出了平衡均方误差 (BMSE),作为衡量学习倾向分数的衡量标准,并基于此提出了 IPS-V2,其有更大的倾向平衡和更小的方差,且不牺牲额外的偏差。

根据以上研究发现,基于 IPS 的去偏方法能够直接针对观测数据中的选择偏差问题,通过分配观测到的不同事件不同的权重来缓解偏差。但是这些方法依赖准确的倾向性得分,如果倾向性得分估计不准确,整个去偏过程可能会失败。

## 6.2 基于双重鲁棒性的去偏方法

文献[109]提出的双重鲁棒性 DR 方法,解决了 IPS 方法在估计无偏损失函数时高方差的缺陷,但需要保证插值模型的准确性,否则会增加预测的偏差。作者联合学习预测模型和插值模型,使得两个模型相互正则化,以提高预测模型和插值模型的准确性:

$$\mathcal{L}_{\text{DR}}(\hat{\mathbf{R}}) = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \hat{e}_{u,i} + \frac{o_{u,i}(e_{u,i} - \hat{e}_{u,i})}{\hat{p}_{u,i}} \quad (19)$$

其中,  $e_{u,i}$  表示预测误差,  $\hat{e}_{u,i}$  表示插值误差。但如果 DR 方法引入的插值模型估计不准确时,会产生比 IPS 方法更大的方差。并且现有方法使用的是模型

无关的方法,并不能灵活改变模型相关的目标。文献[110]将插值模型的目标设为减少 DR 估计的方差这一目标,减少偏差的同时保证双重鲁棒性,采用一种新的双模型学习方法解决插值模型的高偏差问题。

通过对 DR 方法的偏差、方差和泛化边界进行理论分析,Dai 等人<sup>[111]</sup>发现现有的 DR 方法可能由于倾向评分估计不准确和插值误差而导致泛化较差。作者提出了一个通用的学习框架,统一了现有的 DR 方法,而且还开发了一系列适应不同场景的去偏方法。作者提出了两种新的 DR 方法,即 DR-BIAS 和 DR-MSE。DR-BIAS 直接控制 DR 损失的偏置,DR-MSE 灵活平衡偏置和方差,实现更好的泛化性能。Xiao 等人<sup>[112]</sup>改进了 DR 算法,使之能在具有混淆偏差的排名指标进行无偏的评估和学习,并从理论上证明了该估计器具有较小的偏差和方差。作者还提出了一种深度变分信息瓶颈 (IB) 方法,以利用倾向得分的充分性进行估计调整和更好的泛化。Zhang 等人<sup>[113]</sup>提出了 CVR 场景下的多任务逆倾向加权估计法 (Multi-IPW) 和多任务双稳健估计法 (Multi-DR)。其中,多任务逆倾向加权估计用于评估 CVR 任务,多任务双稳健估计用于评估 CTR 任务,插值模型用于加强多任务双稳健估计器。Li 等人<sup>[114]</sup>发现 DR 很难获得准确的插补误差,因为插补模型是用观察到的点击数据学习的,如果可观测到的数据和不可观测到的数据分布差异不同,很容易导致更大的偏差;另外,收集到的数据包含着很多偏差,会导致对倾向的不准确估计。提出了一种多重鲁棒估计器,可以利用多种候选插补模型和倾向估计器实现无偏估计。文献[115]提出 DR 对所有用户-项目对进行插补,可能会导致有害的插补。插补模型不可避免的会对某些用户项目对产生不准确的估计,当插补值与事实严重不符,会对去偏过程产生负面影响。为了解决这个问题,Song 等人提出了保守双稳健策略 (CDR),该策略通过仔细检查输入的均值和方差来过滤输入。理论分析表明,CDR 提供了减少的方差和改进的尾界。

根据以上研究发现,基于 DR 的去偏方法结合了倾向性得分和插值模型,可以从两方面减轻偏差。即使其中一个模型不完全准确时,仍然可以提供一致估计,具有更高的鲁棒性。与此同时,DR 方法增加模型的复杂性。图 8 展示了主要的基于 IPS 和 DR 的去偏方法以及它们之间的演化方向。

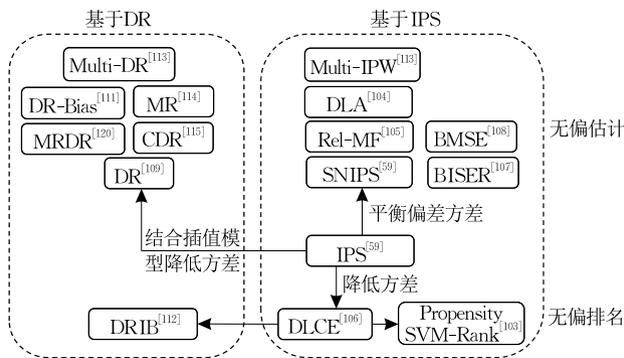


图 8 基于 IPS 和 DR 的方法统计

### 6.3 其他方法

准确的离线评估新模型的潜在收益提升十分重要，离线的估计方法被称为反事实或者离线策略估计方法。上限重要性采样或归一化重要性采样等传统的反事实估计方法，由于没有很好的平衡方差，导致性能欠佳。为了克服这个问题，Gilotte 等人<sup>[116]</sup>建模估计量的偏差，而不是对偏差进行限制，并提出了新的估计方法 NCIS。NCIS 是一种实际可用的推荐模型离线评估器，通过模拟在线实验来快速迭代原型。作者提出了两个上限重要性采样的两个变种，避免基础的重要性采样问题与双重鲁棒性方法高方差的风险。

基于用户交互行为优化排序系统是一个非常重要的研究话题。Oosterhuis 等人<sup>[117]</sup>提出了一种新的干预感知估计器 LTR，该估计器通过考虑在线干预的影响，并结合日志策略的行为和在线干预来纠正位置偏差、信任偏差和项目选择偏差，提高了在线和反事实场景下的学习效率。文献<sup>[118]</sup>发现逆概率加

权 IPS 会破坏 SUTVA 假设，然后提出了一种新的反事实估计器 RIPS，用于评估具有顺序奖励交互的推荐列表。该方法通过模拟实验和在线推荐系统的实时测试，验证了 RIPS 在减少偏差和提高数据效率方面优于现有方法。

解决推荐系统中的 MNAR 问题的现有方法评估时只解决了正样本的偏差，没有考虑负样本的偏差，导致模型的解不是最优的。假设观测到用户点击了某个物品的曝光的概率，是由物品是否被曝光和物品是否被用户喜欢的概率的乘积决定。基于这一假设，文献<sup>[119]</sup>对正样本和负样本使用两个偏好分数，带入到理想损失中修正，这样得到的理想损失在正样本和负样本数据上都是无偏的。

综上所述，推荐系统评估阶段的去偏研究主要包含两种因果策略。其中，基于倾向性评分的方法过度依赖倾向性得分的学习，容易产生较大的偏差和方差；相对于基于倾向性评分的方法，基于双重鲁棒性的方法具有更高的鲁棒性，但是其性能受到插值模型的影响。在评估阶段的去偏方法中，因果推断带来推荐结果的可解释性主要体现在理论分析方面，基于 IPS 和 DR 推导出新的估计器来减少偏差。例如，DLCE<sup>[106]</sup>先用因果图解释偏差的来源，展示了推荐和潜在结果之间的因果关系，然后提供了无偏估计器的理论分析，为用户提供了透明和可信的推荐结果。

表 6 总结和分析了从评估阶段去偏的因果推荐方法。最后，本文归纳了数据、算法和评估阶段所有因果去偏策略的优点和缺点，如表 7 所示。

表 6 评估阶段的因果推荐去偏方法

因果策略	去偏方法	推荐框架	应用场景	性能			文献来源	数据集
				鲁棒性	泛化性	解释性		
基于 IPS	IPS <sup>[59]</sup>	MF	选择偏差	✓			ICML 2016	Yahoo! R3/Coat
	SNIPS <sup>[59]</sup>	MF	选择偏差	✓	✓		ICML 2016	Yahoo! R3/Coat
	Propensity SVM-Rank <sup>[103]</sup>	SVM	位置偏差	✓	✓		WSDM 2017	Yahoo
	DLA <sup>[104]</sup>	DNN	点击偏差	✓	✓		SIGIR 2018	Yahoo! LETOR
	Rel-MF <sup>[105]</sup>	MF	曝光偏差	✓	✓		WSDM 2020	ML-100K/Yahoo! R3
	DLCE <sup>[106]</sup>	MF	曝光偏差	✓	✓	✓	RecSys 2020	Dunnhumby
	BISER <sup>[107]</sup>	MF/AE	曝光偏差	✓	✓		SIGIR 2022	MovieLens/Yahoo! R3/Coat/CiteULike
	BMSE <sup>[108]</sup>	MF	曝光偏差	✓	✓		ICML 2023	Yahoo! R3/Coat/Product
	DR <sup>[109]</sup>	MF	选择偏差	✓	✓		ICML 2019	Yahoo! R3/Coat
	MRDR <sup>[120]</sup>	MF	选择偏差	✓	✓	✓	SIGIR 2021	Yahoo! R3/Coat/ML-100K
基于 DR	DR-BIAS/DR-MSE <sup>[111]</sup>	DCN	选择偏差	✓	✓		SIGKDD2022	Yahoo! R3/Coat/Product
	DRIB <sup>[112]</sup>	MF	选择偏差	✓	✓		WSDM 2022	Dunnhumby/IHDP
	Multi-IPW/Multi-DR <sup>[113]</sup>	MLP	选择偏差	✓	✓	✓	WWW 2020	Ali-CCP
	MR <sup>[114]</sup>	MF/NCf	选择偏差	✓	✓		AAAI 2023	Yahoo! R3/Coat/ML-100K
	CDR <sup>[115]</sup>	MF	选择偏差	✓	✓		CIKM 2023	Yahoo! R3/Coat/KuaiRand
其他方法	NCIS <sup>[116]</sup>	—	离线评估偏差	✓	✓		WSDM 2018	39 online A/B tests
	LTR <sup>[117]</sup>	—	在线评估偏差	✓	✓		WSDM 2021	Yahoo
	RIPS <sup>[118]</sup>	—	位置偏差	✓	✓		SIGKDD2020	Simulation/Real world dataset
	Dual <sup>[119]</sup>	MF	点击偏差	✓	✓		SIGIR 2021	Yahoo! R3/Coat/ML-100K

表 7 各阶段因果去偏策略分析

去偏阶段	因果去偏策略	优点	缺点
数据阶段的去偏	构造反事实空间	有助于揭示和纠正数据中的潜在偏差, 有利于深入理解用户行为	依赖模型假设的准确性, 且计算成本较高
	使用无偏数据	能够直接高效提高预测的准确性, 减少对复杂模型和算法的依赖	无偏数据量小且获取成本极高
算法阶段的去偏	基于因果表示学习	提取潜在的因果特征, 增强模型的鲁棒性和泛化性	需要大量计算资源, 依赖数据集的规模和质量
	基于因果干预	直接对变量进行干预, 有助于更直接地理解因果关系	理论性强, 干预的实施有难度
	基于反事实推理	具有强大的解释性, 支持复杂决策问题	依赖因果假设和因果模型, 面临潜在的可计算性问题
评估阶段的去偏	基于 IPS	直接面向选择偏差问题, 简化模型要求	依赖准确的倾向性得分, 容易产生较大的偏差和方差
	基于 DR	比 IPS 方法更具有鲁棒性, 结合两种模型减轻偏差	复杂度高, 依赖模型的选择和估计的准确性

## 7 开放性研究问题与展望

结合因果推断技术的推荐系统在缓解偏差问题上已经取得了一些成果, 但是对因果推荐的研究还处于一个比较早期的阶段. 目前还有许多方向值得进一步研究以推进本领域的进展, 下面本文总结了以下五个有趣且值得探索的问题.

### (1) 基于因果发现的去偏推荐方法

发现潜在的因果关系是因果推荐的首要工作, 是推荐系统去偏任务的重要研究基础. 现有的基于因果推断的推荐方法, 在进行变量之间的因果推理时所构建的因果图还存在以下问题: 首先, 目前大多数的因果图都比较简单, 对推荐场景中的变量及其因果关系考虑不够全面, 比如模型 DICE<sup>[70]</sup> 和 PDA<sup>[78]</sup> 分别只涉及了用户一致性和项目流行度. 其次, 现有工作对变量之间的因果关系构建不够准确, 一方面是忽略了真正原因的隐藏变量, 另一方面并不能保证满足因果理论的基本假设. 预定义这些因果图需要专业的知识, 但是在现实复杂的环境中, 影响用户决策过程的诸多因素可能超出了专家的理解范围. 因此, 从推荐数据中学习和发现因果机制是一个重要的研究方向. 常用的因果发现有基于条件约束的方法<sup>[121-122]</sup>、基于函数的因果模型方法<sup>[123-124]</sup>等, 新兴的方法比如因果表示学习<sup>[125]</sup>能够发现隐变量和观测变量之间的关系. 另外, 还有相应的因果发现工具比如 causal-learn 提供了最新的因果发现方法的实现, 为研究者们提供了便利. 因此, 如何借助最新因果分析理论和工具, 从观测数据中发现并获取准确的因果关系并估计它们的因果效应, 进一步增强推荐系统的性能, 是一个非常值得探索的研究问题.

### (2) 基于因果的通用性去偏推荐框架

推荐系统中存在着多种多样的偏差, 而目前大多数的去偏方法往往缺乏通用型. 通常情况下, 基于因果图模型的推理去偏框架旨在解决某一种特定类型的偏差或者混杂因素, 忽略了其他可能同时存在的因素, 导致因果推理的过程不够完善. Chen 等人<sup>[64]</sup>从风险差异的角度统一了不同来源的偏差, 提出了一个通用的去偏框架 AutoDebias, 采用元学习自动学习去偏参数, 实现了自适应去偏. 该方法具有较好的去偏通用性和自适应能力, 为研究者们提供了一个很好的去偏思路. 此外, 即使对于解决同一种偏差问题的因果推荐去偏工作, 研究者们也都给出了不同的假设, 它们的因果图也分别包含了不同的变量和不同的因果路径, 这导致同一类型的不同的研究工作之间存在不一致性. 比如, 同样是解决流行度偏差, 文献<sup>[70, 78, 85]</sup>呈现的因果图和变量并不相同. 因此, 如何设计能够全面解决多种类型偏差问题的通用性因果推荐框架, 是一个具有挑战性的问题.

### (3) 基于因果的鲁棒性去偏推荐方法

在实际应用中, 由于存在场景多变、数据动态变化快以及离线评估与在线实验之间存在差距等特性, 许多推荐模型在面对不同的推荐场景、数据和测试环境时, 会出现预测结果不佳的情况, 因此学习更鲁棒和更稳定的推荐模型对学术研究和工业应用都至关重要. 已有的工作表明基于因果性的方法能够识别变量之间的伪相关性, 发现用户行为背后真正的原因, 从而捕获用户更本质的兴趣和偏好, 所以比基于相关性的方法具有更高的鲁棒性和稳定性. 文献<sup>[76]</sup>为了解决推荐系统中训练集和测试集之间数据分布不一致(OOD)问题, 使用因果表示学习来捕捉用户的不变兴趣, 从而缓解偏差问题. 实验验证了

该方法在 OOD 问题上表现出强大的鲁棒性和稳定性。因此,面对推荐系统中未知的偏差、分布偏移的数据等诸多挑战,如何利用因果推断构建具有更高鲁棒性和稳定性的模型非常重要。

#### (4) 基于因果对抗动态环境中的偏差问题

对抗动态环境中的偏差问题是当前推荐系统研究中的一个重要挑战。在现实世界中,用户的行为模式、偏好以及他们与内容互动的方式是不断变化的。这种变化可能由多种因素驱动,包括季节性影响、社会经济因素、新闻事件、流行趋势以及个人生活事件等。所有这些因素都可能导致数据分布的非平稳性,从而对推荐系统的性能产生影响。在这种动态变化的环境中,传统的推荐系统往往会遇到困难,因为它们大多设计为在相对静态的环境中学习固定的用户偏好模式。然而,因果推断技术提供了一个可能的解决框架,允许从变化的数据中识别出真正影响用户行为的因素。这种技术通过模拟不同情况下的用户行为来预测可能的变化,进而适应这些变化。

#### (5) 用于因果去偏推荐的数据集构建

部署在现实世界的推荐系统面临着各种偏差问题。虽然理想情况下在线上与用户一起开发和评估推荐模型能够避免偏差,但由于巨大的成本变得不切实际。离线数据中缺失的大量交互信息是阻碍推荐系统去偏的重要原因。为了解决这个问题,有相关研究工作尝试收集真正的无偏数据集进行公正的无偏评估,其中,比较有名的公开数据集有 Yahoo! R3<sup>[58]</sup>、Coat<sup>[59]</sup> 以及 KuaiRand<sup>[60]</sup> 等。来自华为的研究团队<sup>[62]</sup> 通过在一个大型的广告平台随机展示广告给用户,获得了部分不受推荐机制影响的无偏交互数据,然后通过在线 A/B 测试的方式评估了包括他们方法在内的多个去偏模型。这些包含无偏数据的数据集在基于因果的去偏推荐方法里为无偏的离线评估起到了至关重要的作用,但是如 Yahoo! R3 和 Coat 数据集的规模较小,在许多任务中的使用有所限制。因此,更多的用于去偏推荐的现实世界数据集亟需被构建,这离不开相关研究单位和企业的合作。

## 8 总 结

偏差的存在严重限制了推荐系统的性能,解决这一问题已成为推荐系统领域的重要研究课题。本文从因果推断技术的角度出发,对推荐系统的去偏研究进行了综述。本文根据推荐系统产生偏差的阶段进行分类归纳,分析了在数据阶段、算法阶段和评估阶段不同的因果推断方法在去偏任务上的特点和

性能。同时,对推荐系统的因果去偏研究发展方向进行了探讨和展望。在推荐系统领域,因果推断发挥着至关重要的作用,不仅在增强鲁棒性、公平性和可解释性等方面被广泛研究,还在广告效果评估、智能营销和用户留存分析等业务场景得到应用实践。随着因果推断的理论发展和技术创新,经典方法如潜在因果模型和结构因果模型与深度学习、大模型、图机器学习等人工智能技术的结合正在成为学术界和工业界的研究热点,将推动推荐系统向更智能化和个性化的阶段迈进。

## 参 考 文 献

- [1] Resnick P, Varian H R. Recommender systems. *Communications of the ACM*, 1997, 40(3): 56-58
- [2] Huang Li-Wei, Jiang Bi-Tao, Lv Shou-Ye, et al. Survey on deep learning based recommender systems. *Chinese Journal of Computers*, 2018, 41(7): 1619-1647(in Chinese)  
(黄立威, 江碧涛, 吕守业等. 基于深度学习的推荐系统研究综述. *计算机学报*, 2018, 41(7): 1619-1647)
- [3] Feng Y, Lv F, Shen W, et al. Deep session interest network for click-through rate prediction//*Proceedings of the 28th International Joint Conference on Artificial Intelligence*. Macao, China, 2019: 2301-2307
- [4] Cheng H-T, Koc L, Harmsen J, et al. Wide & deep learning for recommender systems//*Proceedings of the 1st Workshop on Deep Learning for Recommender Systems*. Boston, USA, 2016: 7-10
- [5] Covington P, Adams J, Sargin E. Deep neural networks for Youtube recommendations//*Proceedings of the 10th ACM Conference on Recommender Systems*. Boston, USA, 2016: 191-198
- [6] Chen J, Dong H, Wang X, et al. Bias and Debias in recommender system: A survey and future directions. *ACM Transactions on Information Systems*, 2023, 41(3): 1-39
- [7] Baeza-Yates R. Bias in search and recommender systems//*Proceedings of the 14th ACM Conference on Recommender Systems*. Virtual Event, Brazil, 2020: 2
- [8] Mehrabi N, Morstatter F, Saxena N, et al. A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 2021, 54(6): 1-35
- [9] Abdollahpouri H. Popularity bias in ranking and recommendation//*Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. Honolulu, USA, 2019: 529-530
- [10] Mansoury M, Abdollahpouri H, Pechenizkiy M, et al. Feedback loop and bias amplification in recommender systems //*Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. Virtual Event, Ireland, 2020: 2145-2148

- [11] Peters J, Janzing D, Schölkopf B. *Elements of Causal Inference: Foundations and Learning Algorithms*. Massachusetts Institute of Technology, Cambridge, USA; The MIT Press, 2017
- [12] Pearl J. Causal inference in statistics: An overview. *Statistics Surveys*, 2009, 3: 96-146
- [13] Schölkopf B. *Causality for machine learning//Probabilistic and Causal Inference: The Works of Judea Pearl*. Association for Computing Machinery: New York, USA, 2022: 765-804
- [14] Cai Rui-Chu, Chen Wei, Zhang Kun, Hao Zhi-Feng. A survey on non-temporal series observational data based causal discovery. *Chinese Journal of Computers*, 2017, 40(6): 1470-1490(in Chinese)  
(蔡瑞初, 陈薇, 张坤, 郝志峰. 基于非时序观察数据的因果关系发现综述. *计算机学报*, 2017, 40(6): 1470-1490)
- [15] Luo H, Zhuang F, Xie R, et al. A survey on causal inference for recommendation. *The Innovation*, 2024, 5(2): 1-14
- [16] Wang Y, Liang D, Charlin L, Blei D M. Causal inference for recommender systems//*Proceedings of the 14th ACM Conference on Recommender Systems*. Virtual Event, Brazil, 2020: 426-431
- [17] Gao C, Zheng Y, Wang W, et al. Causal inference in recommender systems: A survey and future directions. *ACM Transactions on Information Systems*, 2024, 42(4): 1-32
- [18] Wu P, Li H, Deng Y, et al. On the opportunity of causal learning in recommendation systems: Foundation, estimation, prediction and challenges//*Proceedings of the 31st International Joint Conference on Artificial Intelligence*. Vienna, Austria, 2022: 5646-5653
- [19] Ricci F, Rokach L, Shapira B. *Introduction to recommender systems handbook*. *Recommender Systems Handbook*. Boston, USA; Springer, 2010: 1-35
- [20] Huang J, Oosterhuis H, de Rijke M. It is different when items are older: Debiasing recommendations when selection bias and user preferences are dynamic//*Proceedings of the 15th ACM International Conference on Web Search and Data Mining*. Arizona, USA, 2022: 381-389
- [21] Vernade C, Cappé O. Learning from missing data using selection bias in movie recommendation//*Proceedings of the 2015 IEEE International Conference on Data Science and Advanced Analytics*. Paris, France, 2015: 1-9
- [22] Jia J, Shang T, Li L, Chen S. De-Biasing user conformity bias and item popularity bias in group recommendation//*Proceedings of the 2023 IEEE 6th Information Technology, Networking, Electronic and Automation Control Conference*. Chongqing, China, 2023: 1559-1563
- [23] Ma Z, Dong Q. Alleviating the unfairness of recommendation by eliminating the conformity bias. *Journal of Physics: Conference Series*, 2021, 2004(1): 1-9
- [24] Yang Y, Li M, Hu X, et al. Exploring exposure bias in recommender systems from causality perspective//*Proceedings of the 2021 IEEE 21st International Conference on Software Quality, Reliability and Security Companion*. Haikou, China, 2021: 425-432
- [25] Khenissi S, Nasraoui O. *Modeling and Counteracting Exposure Bias in Recommender Systems*[M.S. dissertation]. University of Louisville, USA, 2019
- [26] Lerman K, Hogg T. Leveraging position bias to improve peer recommendation. *Public Library of Science One*, 2014, 9(6): e98914
- [27] Collins A, Tkaczyk D, Aizawa A, Beel J. Position bias in recommender systems for digital libraries//*Proceedings of the International Conference on Information*. Cham, Switzerland, 2018: 335-344
- [28] Krauth K, Wang Y, Jordan M I. Breaking feedback loops in recommender systems with causal inference. *arXiv preprint arXiv:2207.01616*, 2022
- [29] Yang L, Cui Y, Xuan Y, et al. Unbiased offline recommender evaluation for missing-not-at-random implicit feedback//*Proceedings of the 12th ACM Conference on Recommender Systems*. Vancouver, Canada, 2018: 279-287
- [30] Saito Y, Joachims T. Counterfactual learning and evaluation for recommender systems: Foundations, implementations, and recent advances//*Proceedings of the 15th ACM Conference on Recommender Systems*. Amsterdam, Netherlands, 2021: 828-830
- [31] Sato M. Online evaluation methods for the causal effect of recommendations//*Proceedings of the 15th ACM Conference on Recommender Systems*. Amsterdam, The Netherlands, 2021: 96-101
- [32] Beel J, Genzmehr M, Langer S, et al. A comparative analysis of offline and online evaluations and discussion of research paper recommender system evaluation//*Proceedings of the International Workshop on Reproducibility and Replication in Recommender Systems Evaluation*. Hong Kong, China, 2013: 7-14
- [33] Zannettou S, Chatzis S, Papadamou K, Sirivianos M. The good, the bad and the bait: Detecting and characterizing clickbait on Youtube//*Proceedings of the 2018 IEEE Security and Privacy Workshops*. San Francisco, USA, 2018: 63-69
- [34] Lin C, Liu X, Xv G, Li H. Mitigating sentiment bias for recommender systems//*Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. Virtual Event, Canada, 2021: 31-40
- [35] Wang N, Qin Z, Wang X, Wang H. Non-clicks mean irrelevant? Propensity ratio scoring as a correction//*Proceedings of the 14th ACM International Conference on Web Search and Data Mining*. Virtual Event, Israel, 2021: 481-489
- [36] Pearl J, Mackenzie D. *The Book of Why: The New Science of Cause and Effect*. New York, USA; Basic Books, Inc., 2018
- [37] Pearl J, Glymour M, Jewell N P. *Causal Inference in Statistics: A Primer*. Chichester, UK; John Wiley & Sons, 2016
- [38] Yao L, Chu Z, Li S, et al. A survey on causal inference. *ACM Transactions on Knowledge Discovery from Data*, 2021, 15(5): 1-46

- [39] Pearl J. *Causality*. Cambridge, UK: Cambridge University Press, 2009
- [40] Haukoos J S, Lewis R J. The propensity score. *The Journal of the American Medical Association*, 2015, 314 (15): 1637-1638
- [41] Glynn A N, Quinn K M. An introduction to the augmented inverse propensity weighted estimator. *Political Analysis*, 2010, 18(1): 36-56
- [42] Bang H, Robins J M. Doubly robust estimation in missing data and causal inference models. *Biometrics*, 2005, 61(4): 962-973
- [43] Funk M J, Westreich D, Wiesen C, et al. Doubly robust estimation of causal effects. *American Journal of Epidemiology*, 2011, 173(7): 761-767
- [44] Agarwal A, Takatsu K, Zaitsev I, Joachims T. A general framework for counterfactual learning-to-rank//*Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. Paris, France, 2019: 5-14
- [45] Mehrotra R, Bhattacharya P, Lalmas M. Inferring the causal impact of new track releases on music recommendation platforms through counterfactual predictions//*Proceedings of the 14th ACM Conference on Recommender Systems*. Virtual Event, Brazil, 2020: 687-691
- [46] Wang Z, Chen X, Wen R, et al. Information theoretic counterfactual learning from missing-not-at-random feedback. *Advances in Neural Information Processing Systems*, 2020, 33: 1854-1864
- [47] Agarwal A, Zaitsev I, Wang X, et al. Estimating position bias without intrusive interventions//*Proceedings of the 12th ACM International Conference on Web Search and Data Mining*. Melbourne, Australia, 2019: 474-482
- [48] Ghazimatin A, Balalau O, Saha Roy R, Weikum G. PRINCE: Provider-side interpretability with counterfactual explanations in recommender systems//*Proceedings of the 13th International Conference on Web Search and Data Mining*. Houston, USA, 2020: 196-204
- [49] Dong Z, Zhu H, Cheng P, et al. Counterfactual learning for recommender system//*Proceedings of the 14th ACM Conference on Recommender Systems*. Virtual Event, Brazil, 2020: 568-569
- [50] Yuan B, Hsia J-Y, Yang M-Y, et al. Improving ad click prediction by considering non-displayed events//*Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. Beijing, China, 2019: 329-338
- [51] Yu J, Zhu H, Chang C-Y, et al. Influence function for unbiased recommendation//*Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. Virtual Event, China, 2020: 1929-1932
- [52] Saito Y. Asymmetric tri-training for debiasing missing-not-at-random explicit feedback//*Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. Virtual Event, China, 2020: 309-318
- [53] Saito K, Ushiku Y, Harada T. Asymmetric tri-training for unsupervised domain adaptation//*Proceedings of the International Conference on Machine Learning*. Sydney, Australia, 2017: 2988-2997
- [54] Yang M, Dai Q, Dong Z, et al. Top-N recommendation with counterfactual user preference simulation//*Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. New York, USA, 2021: 2342-2351
- [55] Zhang S, Yao D, Zhao Z, et al. CauseRec: Counterfactual user sequence synthesis for sequential recommendation//*Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. Virtual Event, Canada, 2021: 367-377
- [56] Wang Z, Zhang J, Xu H, et al. Counterfactual data-augmented sequential recommendation//*Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. Virtual Event, Canada, 2021: 347-356
- [57] Ji J, Li Z, Xu S, et al. Counterfactual collaborative reasoning //*Proceedings of the 16th ACM International Conference on Web Search and Data Mining*. Singapore, 2023: 249-257
- [58] Marlin B M, Zemel R S. Collaborative prediction and ranking with non-random missing data//*Proceedings of the 3rd ACM Conference on Recommender Systems*. New York, USA, 2009: 5-12
- [59] Schnabel T, Swaminathan A, Singh A, et al. Recommendations as treatments: Debiasing learning and evaluation//*Proceedings of the International Conference on Machine Learning*. New York, USA, 2016: 1670-1679
- [60] Gao C, Li S, Zhang Y, et al. KuaiRand: An unbiased sequential recommendation dataset with randomly exposed videos//*Proceedings of the 31st ACM International Conference on Information & Knowledge Management*. Atlanta, USA, 2022: 3953-3957
- [61] Bonner S, Vasile F. Causal embeddings for recommendation //*Proceedings of the 12th ACM Conference on Recommender Systems*. Vancouver, Canada, 2018: 104-112
- [62] Liu D, Cheng P, Dong Z, et al. A general knowledge distillation framework for counterfactual recommendation via uniform data//*Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. Virtual Event, China, 2020: 831-840
- [63] Wang X, Zhang R, Sun Y, Qi J. Combating selection biases in recommender systems with a few unbiased ratings//*Proceedings of the 14th ACM International Conference on Web Search and Data Mining*. Virtual Event, Israel, 2021: 427-435
- [64] Chen J, Dong H, Qiu Y, et al. AutoDebias: Learning to debias for recommendation//*Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. Virtual Event, Canada, 2021: 21-30

- [65] Li H, Xiao Y, Zheng C, Wu P. Balancing unobserved confounding with a few unbiased ratings in debiased recommendations//Proceedings of the ACM Web Conference. Austin, USA, 2023: 1305-1313
- [66] Ding S, Feng F, He X, et al. Interpolative distillation for unifying biased and debiased recommendation//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. Madrid, Spain, 2022: 40-49
- [67] Zhang X, Dai S, Xu J, et al. Counteracting user attention bias in music streaming recommendation via reward modification//Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Washington, USA, 2022: 2504-2514
- [68] Zheng Y, Gao C, Ding J, et al. DVR: Micro-video recommendation optimizing watch-time-gain under duration bias//Proceedings of the 30th ACM International Conference on Multimedia. Lisboa, Portugal, 2022: 334-345
- [69] Zhan R, Pei C, Su Q, et al. Deconfounding duration bias in watch-time prediction for video recommendation//Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Washington, USA, 2022: 4472-4481
- [70] Zheng Y, Gao C, Li X, et al. Disentangling user interest and conformity for recommendation with causal embedding //Proceedings of the Web Conference. Ljubljana, Slovenia, 2021: 2980-2991
- [71] Zhao W, Tang D, Chen X, et al. Disentangled causal embedding with contrastive learning for recommender system //Companion Proceedings of the ACM Web Conference. Austin, USA, 2023: 406-410
- [72] Chen Z, Wu J, Li C, et al. Co-training disentangled domain adaptation network for leveraging popularity bias in recommenders//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, USA, 2022: 60-69
- [73] Xv G, Lin C, Li H, et al. Neutralizing popularity bias in recommendation models//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. Madrid, Spain, 2022: 2623-2628
- [74] Wang S, Chen X, Sheng Q Z, et al. Causal disentangled variational auto-encoder for preference understanding in recommendation//Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval. Taipei, China, 2023: 1874-1878
- [75] Wang W, Lin X, Wang L, et al. Causal disentangled recommendation against user preference shifts. ACM Transactions on Information Systems, 2023, 41(1): 12:1-12:27
- [76] Wang W, Lin X, Feng F, et al. Causal representation learning for out-of-distribution recommendation//Proceedings of the ACM Web Conference 2022. Lyon, France, 2022: 3562-3571
- [77] Wang Z, He Y, Liu J, et al. Invariant preference learning for general debiasing in recommendation//Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Washington, USA, 2022: 1969-1978
- [78] Zhang Y, Feng F, He X, et al. Causal intervention for leveraging popularity bias in recommendation//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, USA, 2021: 11-20
- [79] Wang X, Wang W, Feng F, et al. Causal intervention for fairness in multi-behavior recommendation. arXiv preprint arXiv:2209.04589, 2022
- [80] Wang W, Feng F, He X, et al. Deconfounded recommendation for alleviating bias amplification//Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. Virtual Event, Singapore, 2021: 1717-1725
- [81] He M, Chen X, Hu X, Li C. Causal intervention for sentiment de-biasing in recommendation//Proceedings of the 31st ACM International Conference on Information & Knowledge Management. Atlanta, USA, 2022: 4014-4018
- [82] Wang Y, Guo H, Chen B, et al. Causalint: Causal inspired intervention for multi-scenario recommendation//Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Washington, USA, 2022: 4090-4099
- [83] Xu S, Tan J, Heinecke S, et al. Deconfounded causal collaborative filtering. ACM Transactions on Recommender Systems, 2023, 1(4): 1-25
- [84] He X, Zhang Y, Feng F, et al. Addressing confounding feature issue for causal recommendation. ACM Transactions on Information Systems, 2023, 41(3): 1-23
- [85] Wei T, Feng F, Chen J, et al. Model-agnostic counterfactual reasoning for eliminating popularity bias in recommender system//Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. Virtual Event, Singapore, 2021: 1791-1800
- [86] Mu S, Li Y, Zhao W X, et al. Alleviating spurious correlations in knowledge-aware recommendations through counterfactual generator//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. Madrid, Spain, 2022: 1401-1411
- [87] Wang W, Feng F, He X, et al. Clicks can be cheating: Counterfactual recommendation for mitigating clickbait issue//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. Virtual Event, Canada, 2021: 1288-1297
- [88] He M, Hu X, Li C, et al. Mitigating confounding bias for recommendation via counterfactual inference//Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Grenoble, France, 2022: 524-540
- [89] Xu D, Ruan C, Korpeoglu E, et al. Adversarial counterfactual learning and evaluation for recommender system. Advances in Neural Information Processing Systems, 2020, 33: 13515-13526
- [90] Xiong K, Ye W, Chen X, et al. Counterfactual review-based recommendation//Proceedings of the 30th ACM International Conference on Information & Knowledge Management. Virtual Event, Queensland, Australia, 2021: 2231-2240

- [91] Liu Y, Yen J-N, Yuan B, et al. Practical counterfactual policy learning for Top-K recommendations//Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Washington, USA, 2022; 1141-1151
- [92] Huang W, Zhang L, Wu X. Achieving counterfactual fairness for causal bandit//Proceedings of the AAAI Conference on Artificial Intelligence. Washington, USA, 2022; 6952-6959
- [93] Ge Y, Tan J, Zhu Y, et al. Explainable fairness in recommendation//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, USA, 2022; 681-691
- [94] Lin X, Zhang M, Zhang Y, et al. Fairness-aware group recommendation with Pareto-efficiency//Proceedings of the 11th ACM Conference on Recommender Systems. Como, Italy, 2017; 107-115
- [95] Li Y, Chen H, Xu S, et al. Towards personalized fairness based on causal notion//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, USA, 2021; 1054-1063
- [96] Zhang Q, Zhang X, Liu Y, et al. Debiasing recommendation by learning identifiable latent confounders//Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Long Beach, USA, 2023; 3353-3363
- [97] Ding S, Wu P, Feng F, et al. Addressing unmeasured confounder for recommendation with sensitivity analysis//Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Washington, USA, 2022; 305-315
- [98] Zhu Y, Yi J, Xie J, Chen Z. Deep causal reasoning for recommendations. *ACM Transactions on Intelligent Systems and Technology*, 2024, 15(4): 1-25
- [99] Wang C-D, Deng Z-H, Lai J-H, Philip S Y. Serendipitous recommendation in e-commerce using innovator-based collaborative filtering. *IEEE Transactions on Cybernetics*, 2018, 49(7): 2678-2692
- [100] Wang Z, Zou Y, Dai A, et al. An industrial framework for personalized serendipitous recommendation in e-commerce//Proceedings of the 17th ACM Conference on Recommender Systems. Singapore, 2023; 1015-1018
- [101] Pradel B, Usunier N, Gallinari P. Ranking with non-random missing ratings: Influence of popularity and positivity on evaluation metrics//Proceedings of the 6th ACM Conference on Recommender Systems. Dublin, Ireland, 2012; 147-154
- [102] De Myttenaere A, Grand B L, Golden B, Rossi F. Reducing offline evaluation bias in recommendation systems. *arXiv preprint arXiv:1407.0822*, 2014
- [103] Joachims T, Swaminathan A, Schnabel T. Unbiased learning-to-rank with biased feedback//Proceedings of the 10th ACM International Conference on Web Search and Data Mining. Cambridge, UK, 2017; 781-789
- [104] Ai Q, Bi K, Cheng L, et al. Unbiased learning to rank with unbiased propensity estimation//Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. Ann Arbor, USA, 2018; 385-394
- [105] Saito Y, Yaginuma S, Nishino Y, et al. Unbiased recommender learning from missing-not-at-random implicit feedback//Proceedings of the 13th International Conference on Web Search and Data Mining. Houston, USA, 2020; 501-509
- [106] Sato M, Takemori S, Singh J, Ohkuma T. Unbiased learning for the causal effect of recommendation//Proceedings of the 14th ACM Conference on Recommender Systems. Virtual Event, Brazil, 2020; 378-387
- [107] Lee J-W, Park S, Lee J. Bilateral self-unbiased learning from biased implicit feedback//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. Madrid, Spain, 2022; 29-39
- [108] Li H, Xiao Y, Zheng C, et al. Propensity matters: Measuring and enhancing balancing for recommendation//Proceedings of the 40th International Conference on Machine Learning. Honolulu, USA, 2023; 20182-20194
- [109] Wang X, Zhang R, Sun Y, Qi J. Doubly robust joint learning for recommendation on data missing not at random//Proceedings of the 36th International Conference on Machine Learning. Long Beach, USA, 2019; 6638-6647
- [110] Saito Y. Doubly robust estimator for ranking metrics with post-click conversions//Proceedings of the 14th ACM Conference on Recommender Systems. Virtual Event, Brazil, 2020; 92-100
- [111] Dai Q, Li H, Wu P, et al. A generalized doubly robust learning framework for debiasing post-click conversion rate prediction//Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Washington, USA, 2022; 252-262
- [112] Xiao T, Wang S. Towards unbiased and robust causal ranking for recommender systems//Proceedings of the 15th ACM International Conference on Web Search and Data Mining. Virtual Event, USA, 2022; 1158-1167
- [113] Zhang W, Bao W, Liu X-Y, et al. Large-scale causal approaches to debiasing post-click conversion rate estimation with multi-task learning//Proceedings of the Web Conference. Taipei, China, 2020; 2775-2781
- [114] Li H, Dai Q, Li Y, et al. Multiple robust learning for recommendation//Proceedings of the AAAI Conference on Artificial Intelligence. Vancouver, Canada, 2023; 4417-4425
- [115] Song Z, Chen J, Zhou S, et al. CDR: Conservative doubly robust learning for debiased recommendation//Proceedings of the 32nd ACM International Conference on Information and Knowledge Management. Birmingham, UK, 2023; 2321-2330
- [116] Gilotte A, Calauzènes C, Nedelec T, et al. Offline A/B testing for recommender systems//Proceedings of the 11th ACM International Conference on Web Search and Data Mining. Marina Del Rey, USA, 2018; 198-206

- [117] Oosterhuis H, de Rijke M. Unifying online and counterfactual learning to rank: A novel counterfactual estimator that effectively utilizes online interventions//Proceedings of the 14th ACM International Conference on Web Search and Data Mining. Virtual Event, Israel, 2021: 463-471
- [118] McInerney J, Brost B, Chandar P, et al. Counterfactual evaluation of slate recommendations with sequential reward interactions//Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Virtual Event, USA, 2020: 1779-1788
- [119] Lee J-W, Park S, Lee J. Dual unbiased recommender learning for implicit feedback//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. Virtual Event, Canada, 2021: 1647-1651
- [120] Guo S, Zou L, Liu Y, et al. Enhanced doubly robust learning for debiasing post-click conversion rate estimation//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. Virtual Event, Canada, 2021: 275-284
- [121] Spirtes P, Glymour C N, Scheines R. Causation, Prediction, and Search. Cambridge, Massachusetts, USA: MIT Press, 2001
- [122] Spirtes P L, Meek C, Richardson T S. Causal inference in the presence of latent variables and selection bias//Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence. San Francisco, USA, 1995: 499-506
- [123] Shimizu S, Hoyer P O, Hyvärinen A, et al. A linear non-Gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, 2006, 7(10): 2003-2030
- [124] Zheng Y, Ng I, Zhang K. On the identifiability of nonlinear ICA: Sparsity and beyond//Proceedings of the 36th International Conference on Neural Information Processing Systems. New Orleans, USA, 2022: 16411-16422
- [125] Schölkopf B, Locatello F, Bauer S, et al. Towards causal representation learning. *Proceedings of the IEEE*, 2021, 109(5): 612-634



**YANG Xin-Xin**, Ph. D. candidate.

Her main research interests include recommendation system and causal inference.

**LIU Zhen**, Ph. D. , professor, Ph. D. supervisor. Her main research interests include data mining, recommendation

system and distributed computing.

**LU Si-Bo**, M. S. candidate. His main research interests include recommendation system and causal inference.

**YUAN Ya-Fan**, M. S. candidate. Her main research interests include recommendation system and causal inference.

**SUN Yong-Qi**, Ph. D. , professor, Ph. D. supervisor. His main research interests include artificial intelligence and data mining.

## Background

Recommender system, as an important information filtering technique, has achieved great success in the era of information overload. However, practical applications of recommender systems are often plagued by various types of bias, such as selection bias, exposure bias and popularity bias, etc. Traditional recommendation methods rely on correlation learning to fit data, which fails to deal with the confounding factors in the observational data and can lead to prediction results that deviate from users' real preferences. This bias problem severely limits the performance of recommender systems, and as a result, research on debiasing recommender systems has garnered significant attention. Causality-based statistical methods offer a solution to the shortcomings of correlation-based methods by learning stable causality between variables. Researchers have gradually applied causal inference techniques to the field of recommender systems, resulting in remarkable progress in addressing the bias problem. Existing

work has shown that recommendation methods based on causal inference can effectively alleviate bias and improve recommendation performance.

Our paper provides a comprehensive overview of related work on addressing bias in recommender systems using causal inference techniques. Firstly, this paper details various biases in recommender systems and categorizes them according to the three stages of recommendation. Secondly, the theoretical basis and main techniques of causal inference are introduced. Next, existing debiasing methods are classified based on different causal methods employed at different recommendation stages. This article systematically summarizes and analyzes these methods. Finally, this paper discusses the open issues and future directions in debiasing recommender systems based on causal inference.

This project was funded by the National Key Research and Development Program of China (2019YFB2102500).