无偏学习排序理论基础与应用的最新进展

Shashank Gupta Philipp Hager Jin Huang 阿姆斯特丹大学 阿姆斯特丹大学 阿姆斯特丹大学 阿姆斯特丹，荷兰 阿姆斯特丹，荷兰 阿姆斯特丹，荷兰 s.gupta2@uva.nl p.k.hager@uva.nl j.huang2@uva.nl

Ali Vardasbi 阿姆斯特丹大学 阿姆斯特丹，荷兰 a.vardasbi@uva.nl

Harrie Oosterhuis 拉德堡德大学，奈梅亨，荷兰 harrie.oosterhuis@ru.nl

arXiv:2305.02914v1 [cs.IR] 4 May 2023

摘要

自无偏学习排序（ULTR）领域诞生以来，该领域一直十分活跃，近年来也取得了多项重要进展。本教程既介绍该领域的核心概念，也概述其理论基础方面的最新进展，并介绍其方法的若干应用。

本教程分为四个部分：首先，概述无偏学习排序方法能够处理的不同类型的偏差。其次，全面探讨用户行为日志排序领域的最新估计技术。第三，综述无偏学习排序在现实应用中的已发表成果。第四，讨论无偏学习排序与排序公平性之间的联系。最后，简要展望无偏排序学习研究及其应用的未来。

本教程旨在帮助对开发新的无偏学习排序方案或将其用于现实应用感兴趣的研究人员和业界从业者。

CCS 概念

• **信息系统 → 学习排序**；查询日志分析；推荐系统。

关键词

学习排序；反事实学习排序

ACM 引用格式：

Shashank Gupta、Philipp Hager、Jin Huang、Ali Vardasbi 和 Harrie Oost- erhuis。2023年。《无偏学习排序基础与应用的最新进展》。载于第*46*届 *ACM* 国际信息检索研究与发展会议论文集（*SIGIR '23*），*2023*年*7*月*23*日至*27*日，台湾台北。ACM，纽约州纽约市，4页。https://doi.org/10.1145/3539618.3594247

允许免费为个人或课堂使用而制作本作品的部分或全部内容的数字副本或纸质副本，前提是不得出于营利或商业目的制作或分发副本，且副本须附有本声明及首页上的完整引文。本作品中第三方组件的版权必须得到尊重。其他用途请联系所有者/作者。*SIGIR '23，2023*年*7*月*23*日至*27*日，台北，台湾© 2023年版权归所有者/作者所有。ACM ISBN 978-1-4503-9408-6/23/07。
https://doi.org/10.1145/3539618.3594247

1 动机

学习排序（LTR）算法是现代搜索和推荐系统的基石。传统上，学习排序算法基于从人工标注的相关性判断中进行监督学习。然而，获取相关性判断成本高昂，而且往往与实际用户偏好不一致 [5, 38]。相比之下，点击数据的收集成本较低，通常也更符合用户意图 [16]。然而，点击数据通常是用户偏好的严重偏差信号 [1, 8, 18]，无偏学习排序（ULTR）领域旨在缓解这一问题 [23]。

以往的教程侧重于向信息检索和推荐系统领域的研究人员及从业者介绍该领域的基础知识 [2, 21, 31]。这些内容在当时很有意义，但无偏学习排序领域已显著成熟，此后也取得了基础性进展。在上一次举办教程时，研究最多的交互偏差是位置偏差 [8, 18, 49]。此后，领域内还研究了新的交互偏差，包括信任偏差 [1, 46]、项目选择偏差 [28]、上下文偏差 [51, 53, 57, 58]，和级联位置偏差 [19, 45]。在偏差校正方面，最常用的方法是逆倾向评分（IPS）。然而，如今已知逆倾向评分无法有效校正所有形式的交互偏差 [25, 46]。因此，研究人员开发了多种新的基础性估计技术，以克服逆倾向评分的局限性，例如仿射校正 [30, 46] 和双重稳健估计 [26]；二者都可视为面向无偏学习排序的逆倾向评分扩展。此外，研究人员还提出了与逆倾向评分在原理上截然不同的估计方法，例如双塔模型 [11, 53, 58] 和基于因果推断的方法 [33, 44, 56]。许多无偏学习排序方法着眼于减轻历史数据集中的偏差，而在线学习排序则通过直接与用户交互来减轻偏差 [27, 40, 55]。近期的一系列研究针对在线和离线场景提出了可应用于任一场景的方法，旨在统一无偏学习排序领域 [29, 30]。

近年来，LTR 在应用方面也取得了显著进展 [1, 4, 13, 51, 58]，，包括公平学习排序 [42, 43, 52]。以往的教程主要关注无偏学习排序的基础，对实际应用的关注有限。虽然 LTR 传统上侧重于相关性排序，但如今它通常】【。

承认仅优化相关性可能会导致不公平问题 [3, 41, 54]。在这方面，我们认为，公平学习排序等相近领域的目标与无偏学习排序的使命一致，即为用户提供公平且无偏的排名。为了扩展到大规模应用，公平学习排序领域的工作依赖于无偏学习排序 [39, 47, 52], ，我们希望本教程能鼓励对该领域开展进一步探索。

鉴于无偏学习排序领域取得的这些重大进展，以及其方法日益广泛的应用，我们认为现在正是对该领域最先进成果进行综述的合适时机。通过本教程，我们旨在惠及学术研究人员和业界从业者，无论他们是有兴趣开发新的无偏学习排序解决方案，还是希望在其应用中加以利用。

2 目标

本教程基于以下两个主要目标:

• 向刚接触该主题的学术界人士或从业者介绍无偏学习排序的基本概念，并激发他们的兴趣。
• 全面概述无偏学习排序基础理论与应用领域的重要最新进展，为该领域的新手和专家提供有益参考。

此外，我们还旨在实现以下其他目标:

• 提供关于用户行为日志排序领域数学基础的最新阐释，涵盖可纠正和不可纠正的各种偏差，以及最新的估计技术。由此，这应为用户行为日志排序领域的研究人员开展未来研究奠定坚实基础。

• 深入综述无偏学习排序的现实应用，让从业者能够对应用无偏学习排序的潜在影响形成切合实际的预期。

• 推动并激发跨学科研究，让其他领域的研究人员了解无偏学习排序如何为他们提供帮助。尤其是，我们将重点介绍其与排序公平性主题的联系。

3 对信息检索领域的相关性

近几年，无偏学习排序领域发展迅速，涌现了多项基础性成果，并在信息检索中得到多样化应用。

该领域最早的教程由 Joachims 和 Swaminathan [17], 撰写，他们介绍了搜索和推荐场景中的反事实学习。近期关于反事实评估与学习的教程主要聚焦于基于老虎机反馈的数据 [35, 36]。在推荐系统领域，Chen 等人 [6] 介绍了偏差及其消除策略。

具体而言，针对学习排序，已有教程介绍了离线无偏学习排序 [2, 21, 31] 和在线学习排序 [10]。然而，据我们所知，尚无现有教程涵盖过去三年用户行为日志排序领域取得的重要进展，也未涉及其近期应用，包括公平学习排序 [42, 52]。

4 格式与详细日程

本教程共计三小时（不含休息时间），安排如下:

预备知识（20 分钟）第一场聚焦于预备知识；我们将讨论有监督学习排序的基础，以及位置偏差和反事实学习排序领域的一些早期研究。

• **学习排序基础（5 分钟）**：通过介绍逐点、成对和逐列表学习排序方法，以及从用户交互中学习的概念，讨论有监督学习排序的基础。

• **位置偏差（5 分钟）**：讨论在对用户点击应用传统学习排序方法时产生的位置偏差 [8]。
• **反事实学习排序（10 分钟）**：介绍反事实学习排序的基础知识，讲解如何利用存在位置偏差的用户反馈数据进行学习 [18]。

偏差（40 分钟）本节将介绍无偏学习排序方法可处理的位置偏差以外的交互偏差类型方面的最新进展。

• **信任偏差（10 分钟）**：讨论信任偏差，即无论项目的相关性如何，用户都可能点击排名靠前的项目，因为他们信任搜索引擎 [1, 46]。

• **项目选择偏差（10 分钟）**：讨论项目选择偏差，即用户只能与固定集合中的 k 项目互动，而排名处于前 k 名之外的项目没有曝光机会 [28]。
• **情境偏差（10 分钟）**：讨论情境偏差，即项目的点击概率会受到其在展示页面上周围项目的影响 [51, 53, 57, 58]。
• **级联位置偏差（10 分钟）**：在级联模型下 [8] ，

一个项目的位置偏差不仅取决于其展示排名（许多无偏学习排序研究都如此假设 [8, 18, 49]），还取决于用户此前查看过的项目的相关性 [19, 45]。因此，级联位置偏差通常是对用户行为更现实的假设，例如，用户找到第一个相关结果后往往会停止搜索 [8]。**新型估计方法（70 分钟）**用于校正用户交互数据中偏差的最常见估计技术是逆倾向评分，这一方法由 Wang 等人 [48] 和 Joachims 等人 [18]在开创性研究中提出。然而，近年来，用户行为日志排序领域在新型估计技术方面取得了重要进展。本环节将按照以下安排详细讨论近期的估计技术:

• **仿射校正方法（10 分钟）**：我们将讨论 Vardasbi 等人提出的仿射校正方法 [46]，这是对逆倾向评分的扩展。他们证明，该方法在校正信任偏差时无效。
• **双重稳健估计（10 分钟）**：尽管在离线老虎机学习文献中广受欢迎 [9, 14, 15, 34, 50],，用于学习排序中位置偏差校正的双重稳健估计方法直到最近才被提出 [26]。我们将讨论这一领域中克服了逆倾向评分一些理论和实践缺陷的基础性贡献 [26]。

• **在线与反事实方法（10 分钟）**：在线学习方法是应对学习排序中偏差的另一类方法 [27, 40, 55], ，其以在线/交互式的方式学习用户偏好，而不是纯粹从离线数据中学习。近期，Oosterhuis 和 de Rijke [30] 认为，

无偏学习排序领域可以通过一种新型在线干预感知反事实估计量，受益于在线学习排序。在线学习也被用于从日志策略中收集额外数据，以降低新排序策略的反事实估计方差 [29]。

• **安全反事实策略学习（10 分钟）**：无偏学习排序依赖基于曝光的逆倾向评分，它能够提供无偏且一致的估计，但往往存在较高的方差。尤其是在点击数据较少时，这种方差可能导致无偏学习排序学到次优的排序行为，进而显著增加损害用户体验的风险。近期，Gupta 等人 [12] 提出了一种风险感知型无偏学习排序方法，该方法采用一种新颖的、基于曝光的风险正则化概念，并具有强有力的理论保证，可安全部署。因此，它避免了学习到的策略劣于线上排序系统的风险。

• **边缘化倾向（10 分钟）**：基于逆倾向评分的估计器要做到无偏，高度依赖全支持假设，即日志策略在新策略所考虑的所有动作上都应具有非零概率 [20]。当动作空间很大时，这一假设可能会构成限制，而这在许多现实应用中十分常见。为应对大规模动作空间，Saito 和 Joachims [37] 提出了边缘重要性采样方法，该方法利用所谓的动作嵌入，将方法推广到大规模动作空间。

• **双塔模型（10 分钟）：**迄今为止，无偏学习排序的研究主要聚焦于识别和开发新型偏差校正方法，而对模型设计的关注有限。最近，随着双塔网络的引入，这一趋势正逐渐发生变化 [11, 53, 58]。我们将讨论双塔系列的校正方法。

• **非倾向性方法（10 分钟）：**无偏学习排序中用于偏差校正的主要估计技术基于逆倾向评分 [1, 18, 24, 46, 48]。尽管逆倾向评分具有诸多优势，但它往往存在高方差问题 [26]。最近，无偏学习排序领域已从主要依赖逆倾向评分，逐步转向其他非逆倾向评分的偏差校正技术，例如 Heckerman 方法等基于因果推断的方法 [32]，以及因果似然分解方法 [56]。

综述应用（20 分钟）许多研究将无偏学习排序方法应用于真实场景。我们将讨论现有研究探索的不同场景、这些场景所需的实际调整，以及报告的性能影响。以下简要概述将涉及的场景：

• **Top-K 排序**：无偏学习排序在 Top-K 排名中的应用 [1, 58]。• **信息流推荐**：用于商品信息流推荐的无偏学习排序，用户交互记录自基于二维网格的用户界面 [51, 58]。• **职位推荐**：使用无偏学习排序的职位推荐 [7]。• **公平学习排序**：用于文章推荐的公平策略学习 [43]。

从无偏学习排序到公平学习排序（20 分钟）传统上，公平学习排序依赖人工相关性判断来学习公平的排名策略 [3, 41]。与支持基于点击数据进行相关性排名学习的论点类似，公平学习排序需要采用点击数据，才能得到广泛应用。本教程这一部分将讨论无偏学习排序在公平策略学习中的应用 [22, 52]。

结论与未来工作（10 分钟）最后，我们将总结本教程讨论的要点，并探讨用户行为日志排序领域现有总体反事实方法的一些重要局限 [25]，以及未来研究中一些有前景的方向。

5 所提供的材料

<escape>本教程幻灯片、附有注释的参考文献汇编、相关研究的开放参考代码，以及本次演讲的公开视频录像均可在 tutorialwebsite获取。</escape>

参考文献

[1] Aman Agarwal、Xuanhui Wang、Cheng Li、Michael Bendersky 和 Marc Najork。2019年。解决无偏学习排序中的信任偏差。载于万维网会议。4–14。[2] Qingyao Ai、Jiaxin Mao、Yiqun Liu 和 W Bruce Croft。2018年。无偏学习排序：理论与实践。载于第*27*届*ACM*信息与知识管理国际会议论文集。2305–2306。[3] Asia J Biega、Krishna P Gummadi 和 Gerhard Weikum。2018年。关注度公平性：在排名中摊销个体公平性。载于第*41*届国际*ACM SIGIR*信息检索研究与发展会议。405–414。[4] Adam Block、Rahul Kidambi、Daniel N Hill、Thorsten Joachims 和 Inderjit S Dhillon。2022年。用于效用最大化查询自动补全的反事实学习排序。*arXiv*预印本 *arXiv:2204.10936*（2022年）。[5] Olivier Chapelle 和 Yi Chang。2011年。Yahoo!学习排序挑战赛概览。载于学习排序挑战赛论文集。PMLR，1–24。[6] Jiawei Chen、Xiang Wang、Fuli Feng 和 Xiangnan He。2021年。推荐系统中的偏差问题与解决方案：RecSys 2021教程。载于第*15*届*ACM*推荐系统会议论文集。825–827。[7] Ruey-Cheng Chen、Qingyao Ai、Gaya Jayasinghe 和 W Bruce Croft。2019年。纠正职位推荐中的时效性偏差。载于第*28*届*ACM*信息与知识管理国际会议论文集。2185–2188。[8] Nick Craswell、Onno Zoeter、Michael Taylor 和 Bill Ramsey。2008年。点击位置偏差模型的实验比较。载于*2008*年国际网络搜索与数据挖掘会议论文集。87–94。[9] Miroslav Dudík、John Langford 和 Lihong Li。2011年。双重稳健策略评估与学习。*arXiv*预印本 *arXiv:1103.4601*（2011年）。[10] Artem Grotov 和 Maarten De Rijke。2016年。用于信息检索的在线学习排序：SIGIR 2016教程。载于第*39*届国际*ACM SIGIR*信息检索研究与发展会议论文集。1215–1218。[11] Huifeng Guo、Jinkai Yu、Qing Liu、Ruiming Tang 和 Yuzhou Zhang。2019年。PAL：用于实时推荐系统点击率预测的位置偏差感知学习框架。载于第*13*届*ACM*推荐系统会议论文集。452–456。[12] Shashank Gupta、Harrie Oosterhuis 和 Maarten de Rijke。2023年。通过基于曝光的风险最小化安全部署反事实学习排序。载于*SIGIR 2023*：第*46*届国际*ACM SIGIR*信息检索研究与发展会议。ACM。[13] Ziniu Hu、Yang Wang、Qu Peng 和 Hang Li。2019年。无偏LambdaMART：一种无偏成对学习排序算法。载于万维网会议。2830–2836。[14] Jiawei Huang 和 Nan Jiang。2020年。从重要性采样到双重稳健策略梯度。载于国际机器学习会议。PMLR，4434–4443。[15] Nan Jiang 和 Lihong Li。2016年。用于强化学习的双重稳健离策略价值评估。载于国际机器学习会议。PMLR，652–661。[16] Thorsten Joachims。2002年。使用点击数据优化搜索引擎。载于第八届*ACM SIGKDD*知识发现与数据挖掘国际会议论文集。133–142。[17] Thorsten Joachims 和 Adith Swaminathan。2016年。搜索、推荐与广告投放的反事实评估与学习。载于第*39*届国际*ACM SIGIR*信息检索研究与发展会议论文集。1199–1201。[18] Thorsten Joachims、Adith Swaminathan 和 Tobias Schnabel。2017年。基于偏差反馈的无偏学习排序。载于第十届*ACM*网络搜索与数据挖掘国际会议论文集。781–789。[19] Haruka Kiyohara、Yuta Saito、Tatsuya Matsuhiro、Yusuke Narita、Nobuyuki Shimizu 和 Yasuo Yamamoto。2022年。级联行为模型下排序策略的双重稳健离策略评估。载于第十五届*ACM*网络搜索与数据挖掘国际会议论文集。487–497。

[20] Lihong Li、Wei Chu、John Langford 和 Xuanhui Wang。2011年。基于上下文老虎机的新闻文章推荐算法的无偏离线评估。载于第四届*ACM*网络搜索与数据挖掘国际会议论文集。297–306。[21] Claudio Lucchese、Franco Maria Nardini、Rama Kumar Pasumarthi、Sebastian Bruch、Michael Bendersky、Xuanhui Wang、Harrie Oosterhuis、Rolf Jagerman 和 Maarten de Rijke。2019年。学习排序的理论与实践：从梯度提升到神经网络和无偏学习。载于第*42*届国际*ACM SIGIR*信息检索研究与发展会议论文集。1419–1420。[22] Marco Morik、Ashudeep Singh、Jessica Hong 和 Thorsten Joachims。2020年。控制动态学习排序中的公平性与偏差。载于第*43*届国际*ACM SIGIR*信息检索研究与发展会议论文集。429–438。[23] Harrie Oosterhuis。2020年。从用户与排名的交互中学习：领域的统一。博士学位论文。阿姆斯特丹大学信息学研究所。[24] Harrie Oosterhuis。2022年。根据位置偏置点击反馈，通过双重稳健估计实现无偏学习排序。*arXiv*预印本 *arXiv:2203.17118*（2022年）。[25] Harrie Oosterhuis。2022年。无偏性的终结：揭示基于点击的学习排序的隐含局限。载于*2022*年*ACM SIGIR*信息检索理论国际会议论文集。*ACM*。[26] Harrie Oosterhuis。2023年。通过双重稳健估计纠正无偏学习排序点击反馈中的位置偏差。*ACM*信息系统汇刊41，3（2023年），1–33。[27] Harrie Oosterhuis 和 Maarten de Rijke。2018年。可微无偏在线学习排序。载于第*27*届*ACM*信息与知识管理国际会议论文集。1293–1302。[28] Harrie Oosterhuis 和 Maarten de Rijke。2020年。面向策略的无偏学习排序：针对Top-k 排名。载于第*43*届国际*ACM SIGIR*信息检索研究与发展会议论文集。489–498。[29] Harrie Oosterhuis 和 Maarten de Rijke。2020年。将反事实带入在线环境：高效且无偏的在线排名评估。载于*2020*年*ACM SIGIR*信息检索理论国际会议论文集。137–144。[30] Harrie Oosterhuis 和 Maarten de Rijke。2021年。统一在线与反事实学习排序：一种有效利用在线干预的新型反事实估计量。载于第*14*届*ACM*网络搜索与数据挖掘国际会议论文集。463–471。[31] Harrie Oosterhuis、Rolf Jagerman 和 Maarten de Rijke。2020年。无偏学习排序：反事实与在线方法。载于*Web*会议*2020*伴随论文集。299–300。[32] Zohreh Ovaisi、Ragib Ahsan、Yifan Zhang、Kathryn Vasilaky 和 Elena Zheleva。2020年。纠正学习排序系统中的选择偏差。载于*Web*会议*2020*论文集。1863–1873。[33] Zohreh Ovaisi、Kathryn Vasilaky 和 Elena Zheleva。2021年。离线学习排序系统中的倾向无关偏差恢复。载于第*44*届国际*ACM SIGIR*信息检索研究与发展会议论文集。1763–1767。[34] Yuta Saito。2020年。带有点击后转化的排序指标双重稳健估计量。载于第*14*届*ACM*推荐系统会议论文集。92–100。[35] Yuta Saito 和 Thorsten Joachims。2021年。推荐系统的反事实学习与评估：基础、实现与最新进展。载于第十五届*ACM*推荐系统会议。828–830。[36] Yuta Saito 和 Thorsten Joachims。2022年。交互式系统的反事实评估与学习：基础、实现与最新进展。载于第*28*届*ACM SIGKDD*知识发现与数据挖掘会议论文集。4824–4825。[37] Yuta Saito 和 Thorsten Joachims。2022年。通过嵌入实现大规模动作空间的离策略评估。*arXiv*预印本 *arXiv:2202.06317*（2022年）。[38] Mark Sanderson、Monica Lestari Paramita、Paul Clough 和 Evangelos Kanoulas。2010年。用户偏好与评估指标是否一致？载于第*33*届国际*ACM SIGIR*信息检索研究与发展会议论文集。555–562。

[39] Fatemeh Sarvi、Maria Heuss、Mohammad Aliannejadi、Sebastian Schelter 和 Maarten de Rijke。2021年。理解并缓解异常值对公平排序的影响。arXiv 预印本 arXiv:2112.11251（2021年）。[40] Anne Schuth、Harrie Oosterhuis、Shimon Whiteson 和 Maarten de Rijke。2016年。用于在线快速学习排序的多重交错梯度下降。载于第九届 ACM 网络搜索与数据挖掘国际会议论文集。457–466。[41] Ashudeep Singh 和 Thorsten Joachims。2018年。排名中的曝光公平性。载于第24届 ACM SIGKDD 知识发现与数据挖掘国际会议论文集。2219–2228。[42] Ashudeep Singh 和 Thorsten Joachims。2019年。公平排序的策略学习。arXiv 预印本 arXiv:1902.04056（2019年）。[43] Ashudeep Singh、David Kempe 和 Thorsten Joachims。2021年。不确定性下排名的公平性。《神经信息处理系统进展》34（2021年），11896–11908。[44] Mucun Tian、Chu Guo、Vito Claudio Ostuni 和 Zhen Zhu。2020年。使用异质性处理效应估计进行反事实学习排序。ArXiv abs/2007.09798（2020年）。[45] Ali Vardasbi、Maarten de Rijke 和 Ilya Markov。2020年。基于级联模型的倾向性估计，用于反事实学习排序。载于第43届 ACM SIGIR 信息检索研究与发展国际会议论文集。2089–2092。[46] Ali Vardasbi、Harrie Oosterhuis 和 Maarten de Rijke。2020年。当逆倾向评分不起作用时：无偏学习排序的仿射校正。载于第29届 ACM 国际信息与知识管理会议论文集。1475–1484。[47] Ali Vardasbi、Fatemeh Sarvi 和 Maarten de Rijke。2022年。概率置换图搜索：用于排序公平性的黑盒优化。载于第45届 ACM SIGIR 信息检索研究与发展国际会议论文集（西班牙马德里）（SIGIR '22）。美国计算机协会，纽约，NY，美国，715–725。https://doi.org/ 10 1145 3477495 3532045 . / 。Xuanhui Wang、Michael Bendersky[48]、Donald Metzler 和 Marc Najork。2016年。个性化搜索中带选择偏差的学习排序。载于第39届 ACM SIGIR 信息检索研究与发展国际会议论文集。ACM，115–124。[49] Xuanhui Wang、Nadav Golbandi、Michael Bendersky、Donald Metzler 和 Marc Najork。2018年。个性化搜索中无偏学习排序的位置偏差估计。载于第十一届 ACM 网络搜索与数据挖掘国际会议论文集。610–618。[50] Xiaojie Wang、Rui Zhang、Yu Sun 和 Jianzhong Qi。2019年。针对非随机缺失数据中推荐问题的双重稳健联合学习。载于国际机器学习会议。PMLR，6638–6647。[51] Xinwei Wu、Hechang Chen、Jiashu Zhao、Li He、Dawei Yin 和 Yi Chang。2021年。信息流推荐中的无偏学习排序。载于第14届 ACM 网络搜索与数据挖掘国际会议论文集。490–498。[52] Himank Yadav、Zhengxiao Du 和 Thorsten Joachims。2021年。公平且无偏排序函数的策略梯度训练。载于第44届 ACM SIGIR 信息检索研究与发展国际会议论文集。1044–1053。[53] Le Yan、Zhen Qin、Honglei Zhuang、Xuanhui Wang、Mike Bendersky 和 Marc Najork。2022年。重新审视无偏学习排序的双塔模型。（2022年）。[54] Tao Yang 和 Qingyao Ai。2021年。最大化动态学习排序的边际公平性。载于2021年 Web会议论文集（斯洛文尼亚卢布尔雅那）（WWW '21）。美国计算机协会，纽约，NY，美国， 137 145 10 1145 3442381 3449901 – 。https://doi.org/ 。/ 。Yisong Yue[55] 和 Thorsten Joachims。 。2009以对决老虎机问题的方式交互式优化信息检索系统。载于第26届国际机器学习年会论文集。1201–1208。[56] Haiyuan Zhao、Jun Xu、Xiao Zhang、Guohao Cai、Zhenhua Dong 和 Ji-Rong Wen。2022年。通过因果似然分解进行无偏 Top-k 学习排序。arXiv 预印本 arXiv:2204.00815（2022年）。[57] Zhi Zheng、Zhaopeng Qiu、Tong Xu、Xian Wu、Xiangyu Zhao、Enhong Chen 和 Hui Xiong。2022年。CBR：用于用户建模去偏和点击预测的上下文偏差感知推荐。载于 ACM Web会议论文集（2022年）。2268–2276。[58] Honglei Zhuang、Zhen Qin、Xuanhui Wang、Michael Bendersky、Xinyu Qian、Po Hu 和 Dan Chary Chen。2021年。用于点击去偏的跨位置注意力。载于2021年 Web会议论文集。788–797。