

联想2024：一种提高推荐系统公平性和减少偏见的技术方法



SmartMindAI

专注搜索、广告、推荐、大模型和人工智能最新技术，欢迎关注我

已关注

5 人赞同了该文章

Introduction

推荐系统⁺在双边市场中起着重要作用，旨在为客户和生产者提供个性化的高质量内容。算法的目标是弥合消费者需求与可提供产品之间的差距。构建一个更公平的双边平台对可持续发展至关重要，该平台能够理解和优化人类与算法的互动，并减轻人为反馈循环带来的偏见。然而，流行度偏差可能导致长尾内容得不到公平曝光，这种偏见会使受欢迎的内容不受质量限制地接收不相称的推荐，进而加剧数据不平衡，形成“马太效应⁺”。

针对流行度偏见的研究主要集中在传统训练的排名模型输出后的后处理算法，以关注长尾内容。这种策略包括提高不太受欢迎内容的可见性，采用多元化技术，如浏览排名列表，以淘汰与排名较高内容类似的内容。然而，推荐系统的本质是动态的，目前的后处理方法可能无法长期有效地提高长尾内容的内在价值。此外，它们忽视了双边市场中推荐系统运作的真实情况。我们的研究关注解决由系统性流行度偏见引起的不同级别曝光度差距，而消费者公平更关注解决基于敏感属性的平台上用户群之间的不平等。我们鼓励生产者做出努力，以揭示双边市场动态中长尾内容的有益属性。

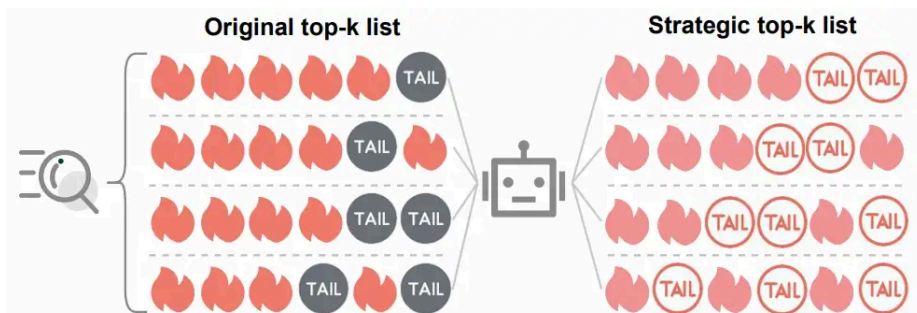


Figure 1: Function of the strategic agent: improve the performance of tail items in top-k recommendation via feature incentives from users' candidate lists.

我们设计了一种基于生产者策略学习代理的排名策略更新方法，以解决长尾内容的优势。每个创作者的目标是最大化自己的曝光，可以修改内容特征以满足受众的最大效用。为了更好地利用内容的具体内容，我们提取尽可能多的语义特征，并将排名算法集中在用户表示与内容特征的内积上。为了指导内容作者的更新，我们提出双优化目标，确保用户前 k 列表中的准确性和公平曝光。然而，这些指标的非可微性质使得梯度下降算法⁺不适合联合优化。为此，我们设计了可微的排名函数，将其整合到端到端训练过程中，并探索这些函数的潜在关系，提出平衡这一权衡的有效方法。

我们通过执行策略代理的指令，通过预见性正则化增强公平曝光。更新后的表示用于计算下一周期的公平性指标，指导当前周期的优化过程。模型能够从倾向于公平性的优化目标中受益。

每轮优化中，关注公平曝光的目标使当前时间训练的用户表示包含更多长尾内容的信息。响应用户表示的动态变化，内容作者策略性地修改内容特征，为特定用户提供最优效用。这些多轮迭代优化

够识别并利用长尾内容的有益特征，促进双边市场的持续发展。

Problem Formalization

我们将问题设定在重排阶段的顶级k推荐中，考虑由用户集合和内容集合组成的平台。

$$\mathcal{U} = \{\mathbf{u}_i \in \mathbb{R}^d\}_{i=1}^m$$

和内容集 $\mathcal{X} = \{\mathbf{x}_j \in \mathbb{R}^d, \|\mathbf{x}\|_2 = 1\}_{j=1}^n$ 组成的平台。内容通过特征向量 $\mathbf{x}_j \in \mathbb{R}^d$ 。其中， d 表示元语义特征的数量。

$\mathcal{C}_i \subseteq \mathcal{X}$ 中生成推荐。

$$\pi_i = \text{rank}(\sigma_i(\mathbf{x}_1), \sigma_i(\mathbf{x}_2), \dots, \sigma_i(\mathbf{x}_c)), \forall i \in [m]$$

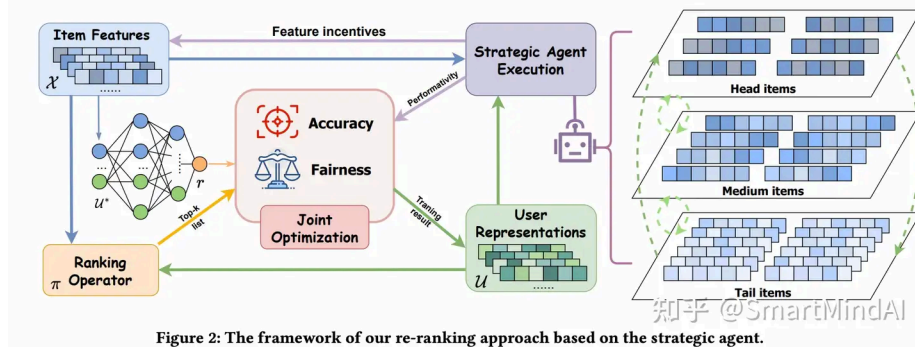
对每个用户*i*，使用个性化的评分函数

$$\sigma_i(\mathbf{x}) = \sigma(\mathbf{x}; \mathbf{u}_i) = \mathbf{u}_i^\top \mathbf{x}$$

进行内容排名，该函数根据用户*i*的分数*c*预测其对候选内容的偏好度。线性评分函数依赖于每个用户*i*的表示向量。

$\mathbf{u}_i \in \mathbb{R}^d$ 系统的目标是学习优秀的 σ 进行重排，其中 $\mathbf{u}_1, \dots, \mathbf{u}_m$ 是通过共同目标学习的表示，以减少偏见。

Methodology



Preliminaries

$\text{DCG}_{@k}(\mathbf{r}_i, \pi_i)$ 是用户推荐列表中前 k 个内容的标准化折扣累积增益 (NDCG) 的计算公式。其中， $\text{DCG}_{@k}$ 表示折扣累积增益， \mathbf{r}_i 表示用户 i 对候选内容的相关性评分， π_i 表示对内容进行排序的运算。NDCG 通过将相关性的收益进行折扣，以反映较低排名内容对总体收益的贡献较少。这种折扣是通过对数函数实现的。

$$\text{NDCG}_{@k}(\mathbf{r}_i, \pi_i) = \frac{1}{\max \text{DCG}_{@k}(\mathbf{r}_i, \pi_i)} \text{DCG}_{@k}(\mathbf{r}_i, \pi_i)$$

其中， $\max \text{DCG}_{@k}$ 是由最高 k 位真实值相关性递减计算得到的 $\text{DCG}_{@k}$ 的最大可能值。内容曝光度通过衡量内容在推荐列表中不同相关性之间的不均衡程度来度量。我们使用基尼系数来度量不同内容在曝光水平上的差异。具体来说，我们按升序对由 π_i 排序的最高 k 位相关性 \mathbf{r}_π 进行排序。

$$r_\pi^1 \leq r_\pi^2 \leq \dots \leq r_\pi^k$$

然后

$$\text{Gini}_{@k}(\mathbf{r}_i, \pi_i) = \frac{\sum_{\ell=1}^k (2\ell - k - 1) r_\pi^\ell}{k \sum_{\ell=1}^k r_\pi^\ell}.$$

知乎

$Gini_{@k}(\mathbf{r}_i, \pi_i)$ 的值越小，推荐列表中不同评分值间的差异越小，反之越大。我们的目标是在保证用户推荐性能的同时，增加尾部产品的曝光度。推荐列表中相关性评分的基尼系数应相对较高，以确保尾部内容在用户候选中得到相对公平的曝光。●动态环境。我们的研究考察了优化过程对用户效益和内容曝光度的影响。每一轮迭代中，通过数据重训练，我们探索了优化的动态性。

$$\mathcal{D}^t = (\mathcal{U}^t, \mathcal{X}^t, \mathcal{C}, \mathcal{U}^*)$$

在训练过程中，我们使用模拟器来生成用户偏好和内容特征之间的相关性作为监督信号。每次训练迭代后，内容作者根据训练得出的用户表示来调整内容特征。 $f(\mathcal{X}^t; \mathcal{U}^t)$ 的战略修改。随后的训练迭代利用新收集的数据来得到

$$\mathcal{D}^{t+1} = (\mathcal{U}^{t+1}, \mathcal{X}^{t+1}, \mathcal{C}, \mathcal{U}^*)$$

动态框架的示意图⁺如图所示。

Differentiable Ranking

本文提出了一种可微排序运算符，用于度量用户效益和内容曝光度。给定用户与候选内容的相关性列表 $\mathbf{r} = [\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_c]^\top$ ，我们的目标是对内容进行排序，并计算前 k 个推荐内容列表的 NDCG 和基尼系数。在重新排序阶段，我们使用连续可微的运算符来近似确定性排序操作，以克服不可微分的限制。

首先，我们定义一个确定性置换矩阵，它是由排列矩阵和一些运算得到的。这个置换矩阵⁺可以用来对列表中的元素进行排序。具体的计算公式如下：

$$P_r[p, q] = \begin{cases} 1 & \text{if } q = \arg \max[(c + 1 - 2p)\mathbf{r} - \mathbf{A}_i \mathbf{1}] \\ 0 & \text{otherwise} \end{cases}$$

其中， p 和 q 表示置换矩阵的行和列， $\mathbf{1}$ 表示所有元素为一的列向量， \mathbf{A}_i 是一个表示列表中元素绝对差的矩阵。这个置换矩阵可以通过使用 softmax 运算符进行连续松弛。

$$\widehat{P}_r[p, :](\tau) = \text{softmax}[(c + 1 - 2p)\mathbf{r} - \mathbf{A}_i \mathbf{1} / \tau]$$

我们使用温度参数 τ 来平滑松弛置换矩阵，使其更接近确定性。较小的 τ 值会使松弛后的置换矩阵更接近确定性，而较大的 τ 值则会使其更接近原始的连续性。我们采用 Sinkhorn 规平 (transformation) 来细化置换矩阵，以获得更准确的逼近结果。

我们使用 $\text{scale}(\widehat{P}_r)$ 来确保矩阵的行和列以概率分布的形式出现，并且总和为一。这有助于减少优化指标在临界点附近梯度的方差，从而提高梯度下降法⁺的效率。通过结合 π 和实际相关性 \mathbf{r} ，我们可以使用 $\text{scale}(\widehat{P}_r)$ 来近似计算当前推荐列表的 NDCG 和基尼系数。这为建立可微的排名操作提供了基础。在关注用户效益的思路中，我们采用了作者提出的方法，其结果为：

$$\begin{aligned} \text{DR-NDCG}_{@k}(\mathbf{r}, \pi)(\tau_1) \\ = \left(\sum_{\ell=1}^k g(r_{(\ell)}) \cdot d(\ell) \right)^{-1} \cdot \sum_{\ell=1}^k [\text{scale}(\widehat{P}) \cdot g(\mathbf{r})]_{\ell} \cdot d(\ell), \end{aligned}$$

我们设计了一个新的可微排名运算符，用于衡量内容曝光度的不平等程度。该运算符使得基于梯度的优化成为可能。具体地，我们定义了符号 $g(x) = 2^x - 1$ 和 $d(x) = \frac{1}{\log_2(x+1)}$ 。在推荐列表中，我们使用符号 \mathbf{r}_{π} 表示内容的实际相关度。

$$\begin{aligned} \text{DR-Gini}_{@k}(\mathbf{r}, \pi)(\tau_2) \\ = \frac{1}{\bar{r}_{\pi} k^2} \sum_{p=1}^k \sum_{q=1}^k |[\text{scale}(\widehat{P}) \cdot \mathbf{r}]_p - [\text{scale}(\widehat{P}) \cdot \mathbf{r}]_q|, \end{aligned}$$

其中 $\bar{r}_{\pi} = \frac{1}{k} \sum_{i=1}^k r_{\pi}^i$ 我们使用变体公式来描述元素之间的差异，以启发对偶排序和分析。这样可以更好地反映我们提出的基尼系数指数的含义。它衡量了候选列表中相关性差异导致的曝光不公。

知乎

考虑到物品的不同受欢迎程度，每个物品所拥有的观众数量也各不相同。我们定义物品 x_j 的效益函数 $s(x_j)$ 为其在目标受众中的排名平均评分。用 \mathcal{U}_j 表示包含物品 x_j 的候选列表的用户集合。然后，我们对针对物品 x_j 的特定用户表示的 l_2 范数⁺进行归一化，以在同一范围内比较不同数量的用户表示。具体而言，我们将 $s(x_j)$ 定义为 $\hat{w}_j^T x_j$ ，其中 \hat{w}_j 是 w_j 的归一化版本，而 w_j 是 \mathcal{U}_j 中所有用户表示的平均值。

接下来，我们定义了系统中策略代理的执行规则。我们假设内容创建者可以根据预测模型的回应来调整物品特征，尽管这会带来一定的成本。根据预定义的成本函数 $cost(x_j, x')$ ，创建者通过最优响应策略来调整物品特征：

$$\Delta_f(x_j) = \operatorname{argmax}_{x': \|x'\|_2=1} [s(x') - \alpha \cdot cost(x_j, x')]$$

其中，

$cost(x, x')$ 表示修改的成本，衡量了执行后的偏离程度，其中 α 是尺度参数⁺。较小的 α 表示更激进的修改策略，导致偏离原始语义信息⁺更大。归一化确保修改后的物品保持 l_2 范数不变，保持维度完整性。公式中的argmax函数是不可微的， Δ_f 依赖于准确性和公平性目标的优化。为了利用预测的引导性，我们提出了一个可微形式。我们使用拉格朗日表达式对公式进行展开，其中 γ 是拉格朗日乘子⁺。

$$L(x', \gamma) = w^T x' - \alpha \|x' - x\|_2^2 + \gamma (\|x'\|_2^2 - 1)$$

随后，通过应用Karush-Kuhn-Tucker (KKT) 条件，我们得到了最终的公式：

$$\Delta_f(x_j) = f(x_j; \hat{w}_j) = \frac{\hat{w}_j + 2\alpha x_j}{\|\hat{w}_j + 2\alpha x_j\|_2}$$

Dynamic Learning and Optimization

Muti-round Dual-objective Optimization.

复排序阶段决定了前 k 推荐列表中的物品展示顺序，对供给方的物品展示产生直接影响。然而，完全追求展示的公平性可能会降低推荐的准确性。因此，在动态交互的每一轮迭代中，我们使用双目标优化来对后处理复排序算法⁺进行优化。这种方法既保持了用户前 k 推荐列表的高效性，又提高了长尾产品的可见度。我们的可微联合优化目标定义如下：

$$\max_{u_i \in \mathcal{U}} \frac{1}{m} \sum_{i=1}^m [\text{NDCG}_{@k}(r_i, \pi_i) + \lambda \text{Gini}_{@k}(r_i, \pi_i)]$$

其中， $r_i = \mathcal{S}(\mathcal{C}_i; u_i^*)$ 是用户 i 及其候选列表中物品的相关度的真实度量。

$\pi_i = \text{rank}(\sigma(\mathcal{C}_i; u_i))$ 是一个对物品进行排序的运算，用于为用户 i 选择物品。正则化参数 λ 用于控制物品显示的公平程度。实质上，我们用 u_i^* 表示用户的真实偏好度量，包含了总候选信息：

$$u_i^* = \frac{u_i'}{\|u_i'\|_2}, \text{ where } u_i' = \frac{1}{|\mathcal{C}_i|} \sum_{x \in \mathcal{C}_i} x$$

在单位 l_2 范数的约束下 u_i^* 可以在优化过程的每一轮中用作监督信号。在当前的第 t 轮结束时，我们得到训练表示 \mathcal{U}^t 。策略代理执行了内容特征的修改：

$$\mathcal{X}^{t+1} = f(\mathcal{X}^t; \mathcal{U}^t)$$

即

$$x_j^{t+1} = \Delta_f(x_j^t)$$

随后，利用基于内容的模拟器 \mathcal{S} 重新计算相关分数，表示为 $r_i^{t+1} = \mathcal{S}(\mathcal{C}_i^{t+1}; u_i^*)$

其中 $\mathcal{C}_i^{t+1} \subseteq \mathcal{X}^{t+1}$ 这样我们就可以使用 r_i^{t+1} 在第 $t+1$ 轮重新训练用户表示。

Agent-based Strategic Learning.

来说，我们利用预测的可行性来实现对公平曝光的前瞻型正则化优化。

$$\max_{\mathbf{u}_i \in \mathcal{U}} \frac{1}{m} \sum_{i=1}^m [\text{NDCG}_{@k}(\mathbf{r}_i, \pi_i) + \lambda \text{Gini}_{@k}(\Delta_f(\mathbf{r}_i), \Delta_f(\pi_i))]$$

至于正则化项，我们用预测形式

$$\Delta_f(\mathcal{C}_i) = \{\Delta_f(x)\}_{x \in \mathcal{C}_i}$$

替换 \mathcal{C}_i ，然后利用 \mathcal{S} 产生的 $\Delta_f(\mathbf{r}_i)$ 替换 \mathbf{r}_i ，即

$$\Delta_f(\mathbf{r}_i) = \mathcal{S}(\Delta_f(\mathcal{C}_i); \mathbf{u}_i^*)$$

同样，排序运算器也被 f 重塑，因为 \mathcal{C}_i 的变化记作

$$\Delta_f(\pi_i) = \text{rank}(\sigma(\Delta_f(\mathcal{C}_i); \mathbf{u}_i))$$

策略代理的响应可以实时反馈给当前的优化过程，为非受欢迎特性提供更多激励。根据公式获得优化损失的直观形式，用于 t -第轮对用户表示 \mathbf{u}_i 进行训练。基于代理策略学习的损失函数如下：

$$\begin{aligned} \mathcal{L}(\mathcal{D}_t) = & -\frac{1}{m} \left[\sum_{i=1}^m \text{DR-NDCG}_{@k}(\mathbf{r}_i, \pi_i)(\tau_1) \right. \\ & \left. + \lambda \sum_{i=1}^m \text{DR-Gini}_{@k}(\Delta_f(\mathbf{r}_i), \Delta_f(\pi_i))(\tau_2) \right]. \end{aligned}$$

公式优化当前用户效用和预期内容展示，可以充分利用策略代理在动态学习过程中的作用。

Algorithm 1 Multi-round Training Process

- 1: Training procedure at time period t ;
- 2: **Input:** Item features \mathcal{X}^t , initialized user representations \mathcal{U} , the relevance simulator \mathcal{S} , hyperparameters λ, α, τ ;
- 3: **Output:** Updated user representations \mathcal{U}^t , modified item features \mathcal{X}^{t+1} ;
- 4: **for** epoch **do**
- 5: Get $\Delta_f(\mathcal{X}^t) = f(\mathcal{X}^t; \mathcal{U}^t)$ based on Eq.(12);
- 6: **for** user $i \in \{1, 2, \dots, m\}$ **do**
- 7: Get $\mathcal{C}_i^t \subseteq \mathcal{X}^t$ and $\Delta_f(\mathcal{C}_i^t) \subseteq \Delta_f(\mathcal{X}^t)$
- 8: Use $\hat{\mathbf{P}}$ based on $\pi_i, \Delta_f(\pi_i)$ to generate the top- k list;
- 9: Use \mathcal{S} to get $\mathbf{r}_i, \Delta_f(\mathbf{r}_i)$;
- 10: Calculate $\text{NDCG}_{@k}(\mathbf{r}_i, \pi_i)$ as Eq.(7)
- 11: Calculate anticipated $\text{Gini}_{@k}(\Delta_f(\mathbf{r}_i), \Delta_f(\pi_i))$ as Eq.(8)
- 12: **end for**
- 13: Calculate total optimization loss as Eq.(15);
- 14: Update \mathcal{U}^t via gradient-descent;
- 15: **end for**
- 16: Get $\mathcal{X}^{t+1} = f(\mathcal{X}^t; \mathcal{U}^t)$ based on Eq.(12)
- 17: **return** $\mathcal{U}^t, \mathcal{X}^{t+1}$

知乎 @SmartMindAI

Experiments

Content-based Simulator

Datasets.

了其中的用户评论数据。工业数据集是针对笔记本消费的用户的行为日志数据。这些数据集提供了我们所需的统计信息，包括用户数量、餐厅数量、评论数量、物品数量和点击次数。

Table 2: Statistics of the Yelp and Industrial dataset.

Data set	Yelp restaurant	Industrial
#Users	1,377 (242)	11,238 (351)
#Items	22,197 (1,103)	1,214 (951)
#Interactions	113,852	270,259
Density	0.00372	0.0198

Design and Training.

我们设计了一个模拟器，用于捕捉用户和内容特征之间的关系。模拟器采用了多层感知器⁺（MLP）架构，具有ReLU激活函数⁺。模拟器包括五个层，每一层的输出维度减半。我们通过聚合用户评价的内容特征来推导出用户的真实特征。在训练过程中，我们使用二元交叉熵⁺损失函数对模拟器进行训练，并在测试集上获得了良好的准确率。

Experimental Results (RQ1&2)

我们在图中展示了通过两个数据集进行的实验结果。我们使用不同的 λ 值来呈现我们的发现，以确保最终的NDCG保持在0.7以上并避免过拟合⁺。具体来说，对于Yelp数据集，我们设置 λ 为5和10；对于工业数据集，我们设置 λ 为2和4。

The Effect of Strategic Agents (RQ1).

首先，我们分析了公平曝光对系统性能的影响。结果显示，系统在处理公平性后，Gini系数始终高于处理前。传统的重新排序方法(MMR)仅基于准确性，无法维持公平曝光。这是因为仅基于准确性的方法导致系统陷入同质化状态，使得NDCG值向1收敛。接着，我们对比了重新训练和不重新训练两种方法。在每一轮迭代中，重新训练的方法获得的NDCG值始终高于不重新训练的方法。这是因为重新训练的模型更准确地捕获了当前优化迭代中的信息。对于有更多候选者的工业场景，重新训练与不重新训练之间的差异在 α 值较小时尤为显著。然后，我们分析了前瞻性正则化对公平曝光优化的影响，对比了基于代理的干预（□）与非代理的干预（△）。在较高的 λ 值条件下，基于代理的方法总能优化Gini系数。这表明，着眼于未来干预的优化策略能够更快地引导系统趋向均衡曝光。尤其是在较低的 λ 值下，基于代理的方法在初始训练轮次中表现更优。结果对比表明，利用预测的反馈性能可以实现中立性偏见目标的公平曝光优化，操作更加有效和高效。最重要的是，基于代理的方法通过提高Gini系数同时改进NDCG值，实现了更好的绩效，这在 λ （↑）变化后的图中得到了说明。这一成果表明，基于代理的方法巧妙地在有限的时间步骤内识别并利用了少受欢迎内容的有利特征，从而实现了在提高长尾内容可见度的同时，同步提升用户效用的平衡。

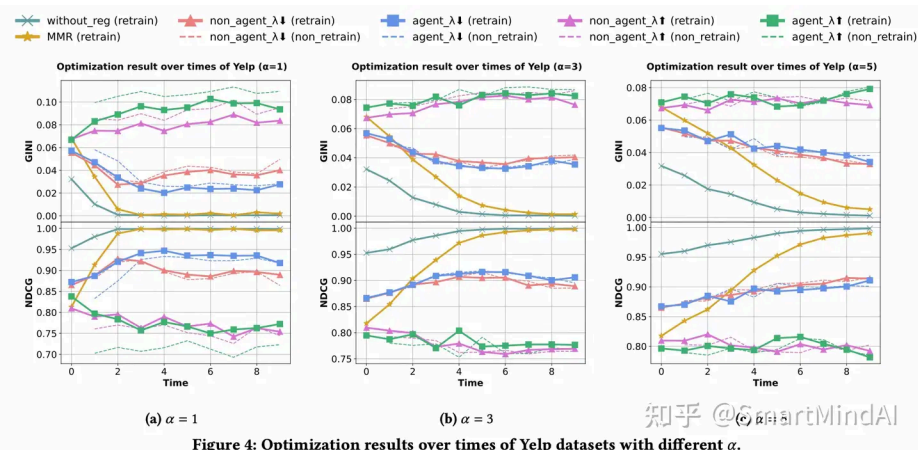


Figure 4: Optimization results over times of Yelp datasets with different α .

知乎

我们探讨了动态设置中正则化参数 λ 和修改成本缩放参数 α 的作用。较大的 λ 值增强了系统公平展示优化的能力，倾向于用较不受欢迎的属性替换初始top- k 列表中的流行属性。这种调整鼓励战略性代理激励长尾内容的特征，导致Gini系数逐步提高，NDCG值下降。较小的 λ 值允许Gini系数减少，加速系统同质化的移动。较高的 α 值意味着修改内容属性的成本更高，减少不同内容类别间特征激励。较大的 α 能更好地平衡系统迭代的演进。参数 α 的值越大，再培训和非再培训方法之间的差异越小。

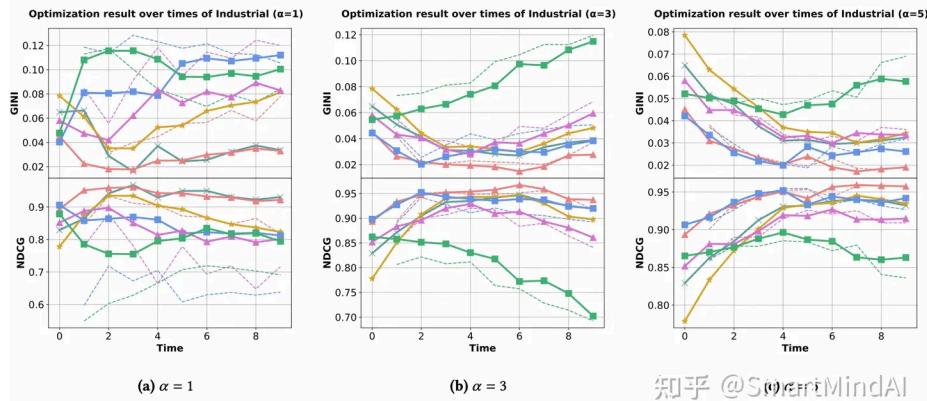


Figure 5: Optimization results over times of industrial datasets with different α .

Effectiveness Results of Debias (RQ3)

我们提出的去偏机制通过策略性重排序，在两个数据集的top- k 推荐列表中展现了有效性。分析显示，在特定时间点，对五种不同受欢迎程度的内容类别的推荐频率进行了调整。与原始的top- k 列表相比，我们的方法降低了热门内容的推荐频率，增加了较不受欢迎内容的曝光。去偏的程度可以通过调整参数灵活控制。例如，较低的 α 值可以提升Yelp中[1-5,6-10]类和工业领域的[1-15,16-30]类内容的曝光率，而较高的 α 值则适度控制了热门内容的推荐，如[20+]和[60+]类，避免它们被过度降级。

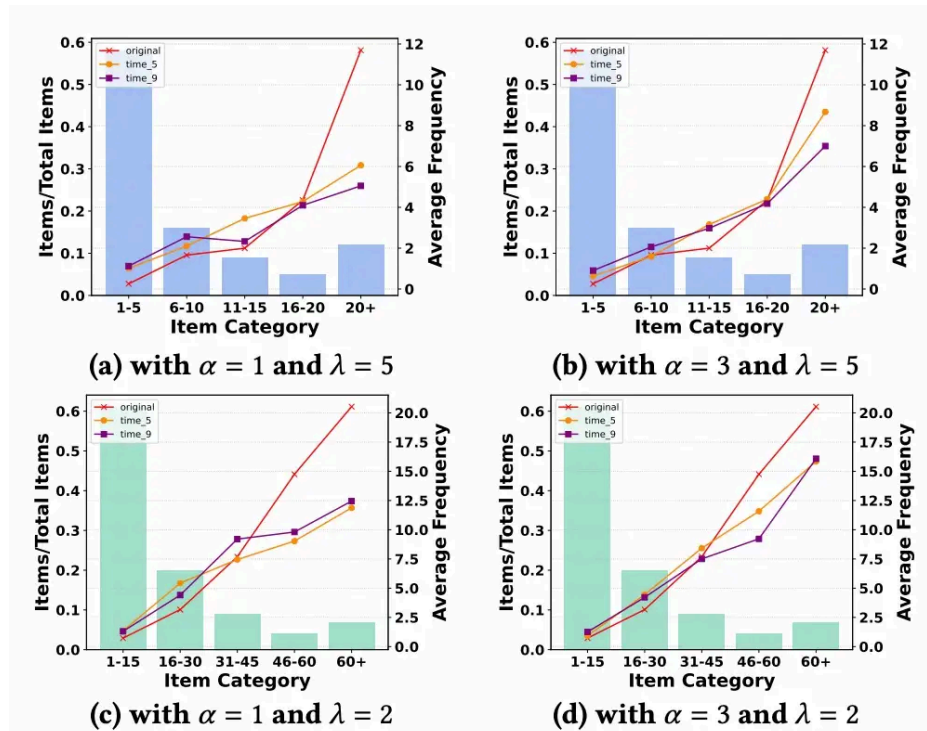


Figure 6: Average frequency of different item categories in two datasets recommended by strategic agents when $k = 10$.

原文《Performative Debias with Fair-exposure Optimization Driven by Strategic Agents in Recommender Systems》

发布于 2024-07-23 10:59 · IP 属地北京