arXiv:2302.03419v1 [cs.IR] 7 Feb 2023

推荐中的准确率-偏差权衡自采样训练与评估

DugangLiu[1,2],Yang Qiao[3], Xing Tang[3], Liang Chen[3], Xiuqiang He[3], Weike Pan[1(✉)]，以及 Zhong Ming[1(✉)]

[1] 深圳大学计算机科学与软件工程学院，中国深圳 dugang.ldg@gmail.com，{panweike,mingz}@szu.edu.cn[2] 广东省人工智能与数字经济实验室（深圳）[3] FIT，腾讯，中国深圳{sunnyqiao,shawntang,leocchen,xiuqianghe}@tencent.com

摘要。 去偏推荐研究已取得令人鼓舞的成果。然而，要将其应用于工业推荐，仍有一些问题需要解决。例如，现有方法大多需要特定的数据、架构和训练方法。本文首先通过一项在线研究指出，在工业推荐中任意消除所有偏差，未必总能带来预期的性能提升。针对无法获取随机化数据集的情况，我们提出一种新颖的自采样训练与评估（SSTE）框架，以实现推荐中的准确率-偏差权衡，即消除有害偏差并保留有益偏差。具体而言，SSTE利用自采样模块，从原始训练数据和验证数据中生成若干具有不同偏差程度的子集。自训练模块基于这些子集推断有益偏差并学习更优的权衡方案；自评估模块则旨在利用这些子集构建更可信的参考，以反映优化模型。最后，我们在两个数据集上进行了大量离线实验，以验证SSTE的有效性。此外，我们将SSTE部署于一款知名理财产品腾讯理财通的首页推荐中，并在在线 A/B 测试中取得了非常理想的结果。

关键词： 去偏推荐 · 自采样 · 自训练 · 自评估。

1 引言

用户在与推荐系统交互的过程中，难免会受到各种偏差的影响，从而导致反馈数据存在固有的可变性。因此，收集到的数据可能无法反映用户的真实偏好 [11]。如果忽视这些偏差，基于反馈数据训练的推荐模型就会继承甚至放大这些偏差的影响，这不利于推荐系统的长期健康发展

。因此，如何合理有效地缓解反馈数据中的偏差问题，是一项重要挑战。

现有的去偏推荐方法主要分为两类：一类使用随机化数据集，另一类则不使用随机化数据集。随机化数据集是通过特定的均匀策略而非推荐模型收集的；由于采用了随机选择操作来分配物品，因此可将其视为良好的无偏代理 [2]。第一类方法利用随机化数据集所包含的无偏信息，重点设计不同的联合训练模块，将这些信息迁移到使用有偏反馈数据训练的推荐模型中 [2,9,3]。如果无法获得随机化数据集，第二类方法则主要通过确保优化目标的无偏性，并引入一些理论框架来指导模型架构的设计 [15,11,19]。尽管现有去偏推荐方法已展现出良好的效果，但它们在工业推荐中的应用仍缺乏充分的见解，因为大多数方法都需要特定的数据、架构和训练方法。

具体而言，一个很少被考虑和回答的重要问题是：在工业推荐中去除所有偏差是否是一个理想目标？为了初步了解这一问题，我们首先在真实推荐场景中开展了一项为期三周的在线研究，部署近似均匀策略，与基准模型进行比较。该推荐场景来自腾讯理财通的首页推荐，这是中国最大的金融推荐场景之一，其展示首页如图 1 左侧所示。请注意，关于该场景和评估指标COPM的更多信息见第 4.4 节。从图 1 右侧可以看出，均匀策略在初期会带来预期的性能提升，这可能是因为随机选择操作带来了意料之外的推荐。然而，在后期，均匀策略的优势未能保持，性能逐渐下降至与基准模型相近。我们认为，这可能是因为均匀策略去除了有助于提升基准模型性能的有益偏差。例如，在这一推荐场景中，高收益基金产品理应获得更多曝光。总的来说，这意味着在工业推荐中任意去除所有偏差，未必能持续带来预期的性能提升。

因此，本文提出采用准确率-偏差权衡，而非消除所有偏差。为实现这一目标，我们进一步提出一种简单而有效的自采样训练与评估（SSTE）框架，在保留有益偏差的同时去除有害偏差。具体而言，我们的 SSTE 包含三个定制模块：1）自采样模块从原始训练集和验证集中生成若干对应的辅助子集，这些子集具有不同程度的偏差；2）自训练模块将原始训练集与辅助子集结合起来进行联合学习，以推断有益偏差并实现更好的准确率-偏差权衡；3）自评估模块将原始验证集与辅助子集结合起来，构建一个更

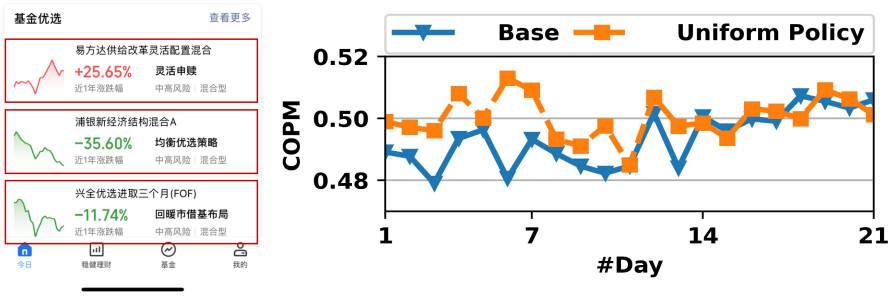


图 1. 腾讯理财通首页的推荐场景，以及开展为期三周的在线研究所得结果。红框标示的内容为推荐物品。请注意，出于保密原因，我们对实际的 COPM 值进行了转换。

合理的参照，以便更好地反映优化模型的离线表现。我们在一个公开数据集和一个真实产品数据集上进行了广泛的离线实验，以验证 SSTE 的有效性，包括无偏评估和兼容性分析。此外，我们还通过在线 A/B 测试展示了 SSTE 的优势。

2 相关工作

本节简要回顾两个研究主题的相关工作，包括去偏推荐和去偏评估。

去偏推荐。根据涉及的反馈数据类型，现有去偏推荐方法主要可分为两类，即基于随机化数据集和不依赖随机化数据集的去偏推荐。前者旨在引入随机化数据集作为无偏代理，随后可以设计多种方法，将其与原始有偏反馈数据进行联合训练，以利用这些无偏信息的指导 [2,9,3]。后者则考虑在无法获取随机化数据集时如何缓解偏差问题。主要技术包括假设并建模特定偏差与某些特征之间的生成机制 [13,12], ，或引入一些理论框架，为该偏差问题构造相应的无偏估计量 [18,15,11,19]。然而，现有方法大多需要特定数据、架构和训练方法，这阻碍了工业推荐场景中对去偏推荐的充分探索和深入理解。我们的 SSTE 旨在弥合这一方向上的差距。

去偏评估。 由于反馈数据存在固有偏差，传统评估指标可能无法反映推荐模型的真实性能，导致离线评估与在线评估之间出现差异。为解决这一问题，先前的大多数研究主要从测量设计和样本设计两个方面入手。前者旨在为传统指标设计相应的无偏版本，或提出

一些新的无偏评估器 [8,20,5], ，而后者则侧重于设计能够构建无偏验证集的方法 [7,2]。然而，由于潜在风险（即评估误差）难以控制，现有的大多数方法不便于应用于工业推荐。与这些方法不同，我们提出的 SSTE 是一种简单有效的自评估方法，其潜在风险可控。

3 所提出的方法框架

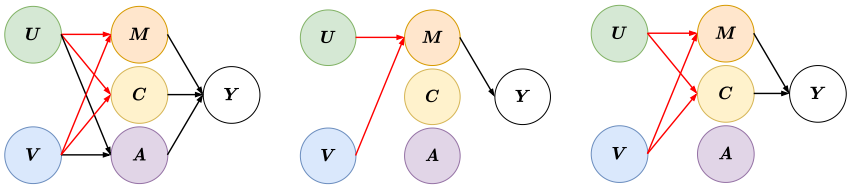
本文重点关注在没有随机化数据集的情况下缓解隐式反馈中的偏差问题。假设训练集 $\mathcal{D}_{tr} = \{(\boldsymbol{x}_i,\ y_i)\}_{i=1}^{m}$由来自潜在分布 $P(\boldsymbol{x},\ y)$的数据构成，其中m为训练实例的数量。 $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_d$是一个 d维特征空间， $\mathcal{Y} = \{0, 1\}$ 是一个标签空间。此外，验证集 $\mathcal{D}_{val} =\{(\boldsymbol{x}_j,\ y_j)\}_{j=1}^{n}$由来自潜在分布 $Q(\boldsymbol{x},\ y)$的数据构成，其中 n为验证实例的数量。请注意， $y = 1$ 和 $y = 0$ 分别表示训练或验证实例为正反馈和负反馈。

3.1 准确率-偏差权衡

由于图 1 的结果表明，在工业推荐中任意移除所有偏差可能并不是理想的选择，因此我们提出了一个新的准确率-偏差权衡目标，以获得更理想的性能提升，其核心思想是将反馈数据中的所有偏差视为有害偏差和有益偏差的组合。为便于理解我们的目标与现有工作的差异，我们分别在图 2 中给出了传统推荐、去偏推荐以及所提出的新目标的因果图。我们使用 U、 V、 M、 C、 A和 Y分别表示用户、物品、真实匹配偏好（即 U对 V的特定偏好）、有益偏差效应（即由有益偏差导致的偏好偏移，例如高收益基金的高曝光偏差）、有害偏差效应（即由有害偏差导致的偏好偏移，例如位置偏差）以及反馈标签。如图 2(a) 和图 2(b) 所示，传统推荐方法会编码有害偏差效应，而去偏推荐方法会同时移除有益偏差效应和有害偏差效应。从图 2(c) 可以看出，与它们不同，我们的目标是在保留有益偏差效应的同时移除有害偏差效应。

3.2 架构

我们提出了一种简单但有效的自采样训练与评估（SSTE）框架，以实现所需的准确率-偏差权衡，其整体架构如图 3 所示。给定训练集 $\mathcal{D}_{tr}$ 和验证集 $\mathcal{D}_{val}$，自采样模块基于 $\mathcal{D}_{tr}$ 和 $\mathcal{D}_{val}$构建一组具有不同偏差程度的辅助子集，即 $\mathcal{A}_{tr} = \{\hat{\mathcal{D}}_{tr}^{i}\}_{i=1}^{T_1}$ 和 $\mathcal{A}_{val} = \{\hat{\mathcal{D}}_{val}^{i}\}_{i=1}^{T_2}$。 $\hat{\mathcal{D}}_{tr}^{i}$ 是根据特定策略从 $\mathcal{D}_{tr}$ 中采样得到的辅助子集，



(a) 传统推荐系统 (b) 去偏推荐系统 (c) 我们的目标

图2. 因果图分别展示了 (a) 传统推荐、(b) 去偏推荐，以及 (c) 所提出的准确率-偏差权衡方案，其中 U、 V、 M、C、 A 和 Y 分别表示用户、物品、真实匹配偏好、有益偏差效应、有害偏差效应和反馈标签。

T_1 是为 $\mathcal{D}_{tr}$ 配备的辅助子集数量。ˆ $\mathcal{D}_{val}^{i}$ 和 T_2 对于 $\mathcal{D}_{val}$ 的定义类似。随后，自训练模块接收 $\mathcal{D}_{tr}$ 和 $\mathcal{A}_{tr}$，并通过与一些共享参数联合训练来更新推荐模型 Θ 。更新后的模型 Θ 随后传递给自评估模块。结合 $\mathcal{D}_{val}$ 和 $\mathcal{A}_{val}$ 后，使用所定义的新评估方法获得当前训练迭代对应的性能。如果未满足收敛条件，则继续交替执行自训练模块和自评估模块。一旦满足收敛条件，便输出优化后的推荐模型 Θ^* 。

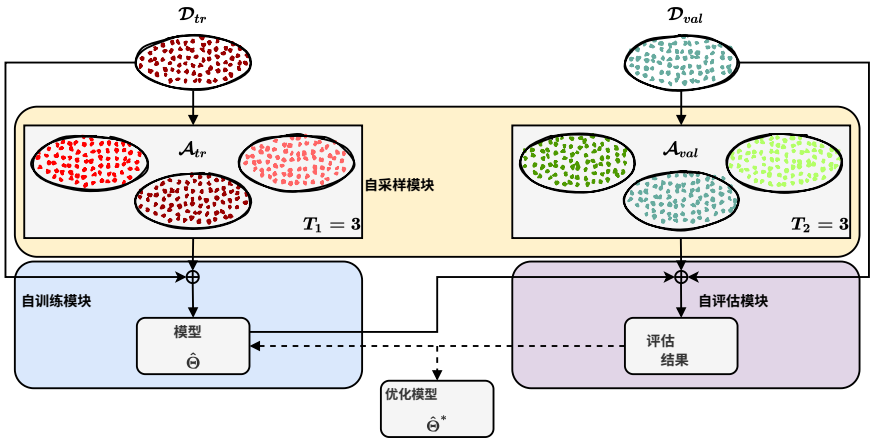


图 3. 所提出的 SSTE 的整体架构，其核心组件包括自采样模块、自训练模块和自评估模块。

3.3 训练

接下来，我们将基于训练过程对各个模块进行详细介绍。

自采样模块。我们提出一种基于截断逆倾向得分（tIPS）的辅助子集采样方法。具体来说，以训练实例为例，我们首先通过去偏推荐中的一些现有 IPS 估计器，获得每个实例 () 的采样概率

，其中较高的 $p(\boldsymbol{x}_i)$ 表示该样本对于去偏更为重要。接下来，我们选取一组不同的截断阈值 $\{\epsilon_{tr}^i\}_{i=1}^{T_1}$，并分别使用每个阈值调整得到的 $p(\boldsymbol{x}_i)$，即当 $p(\boldsymbol{x}_i) < \epsilon_{tr}^i$ 时保留 $p(\boldsymbol{x}_i)$，否则将 $p(\boldsymbol{x}_i)$ 修改为 1。这意味着，我们通过对不同比例的重要样本提供更多保护，来控制去偏程度。最后，根据修改后的采样概率，我们可以得到一组 $\mathcal{D}_{tr}$ 的辅助子集，即 $\mathcal{A}_{tr} = \{\hat{\mathcal{D}}_{tr}^i\}_{i=1}^{T_1}$，其中每个辅助子集对应一个截断阈值。同样，通过为验证实例设置 $\{\epsilon_{val}^i\}_{i=1}^{T_2}$，我们可以得到与 $\mathcal{D}_{val}$ 对应的 $\mathcal{A}_{val} = \{\hat{\mathcal{D}}_{val}^i\}_{i=1}^{T_2}$。这一采样操作的思路是，基于反馈数据本身，模拟具有不同偏差程度和不同偏差类型的辅助子集。请注意，自采样模块可以仅作为预处理操作执行一次，也可以在每轮训练后重新执行。我们在图 4 中给出了采样过程的示例。

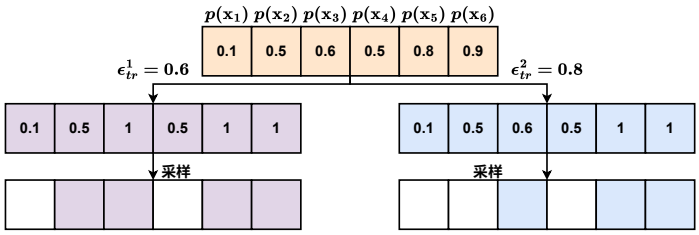


图 4. 采样过程示意图。

自训练模块。为了结合 $\mathcal{D}_{tr}$ 和 $\mathcal{A}_{tr}$ 推断有益的偏差信息，并约束模型以实现更佳的准确率与偏差权衡，我们可以使用任何具有一些共享参数的模型架构进行~~自训练。形式上，设 $\Theta = \{\theta, \theta_s\}$ 为与 $\mathcal{D}_{tr}$相关的模型参数，$\hat{\Theta} = \{\theta, \theta_s\}$ 为与 $\mathcal{A}_{tr}$相关的模型参数，其中 θ_s 为共享参数。我们的 SSTE 的最终优化目标函数可表示如下：

$$\min_{\widetilde{\Theta},\hat{\Theta}} \mathcal{L}_{SSTE} = \mathcal{L}_{\mathcal{D}_{tr}}\left(f\left(\widetilde{\Theta}, \boldsymbol{x}_i\right), y_i\right) + \mathcal{L}_{\mathcal{A}_{tr}}\left(g\left(\hat{\Theta}, \boldsymbol{x}_j\right), y_j\right) + \lambda\|\widetilde{\Theta}\| + \lambda\|\hat{\Theta}\|, \quad (1)$$

其中，$(\boldsymbol{x}_i, y_i) \in \mathcal{D}_{tr}$、$(\boldsymbol{x}_j, y_j) \in \mathcal{A}_{tr}$、$f(\cdot)$ 和 $g(\cdot)$ 是映射函数，而 λ 和 $\|\cdot\|$ 分别是权衡权重和正则化项。请注意，我们将在推理阶段采用模型 Θ。式（1）的直观解释是：在模拟不同环境的数据上进行联合训练，会迫使模型隐式地区分所有偏差；其中，达成更多共识且编码更多信息的偏差更有可能带来益处，因为它们带来的性能提升在不同环境中更为稳健。

ˆ自评估模块。为了更好地离线捕获最优模型 Θ^*，我们提出了一种偏差鲁棒评估方法。具体而言，在第 i 次训练迭代时，模型 $\hat{\Theta}$ 在 $\mathcal{D}_{val}$ 和 $\hat{\mathcal{D}}^i_{val}$ 上的性能分别为 $e(\mathcal{D}_{val})$ 和 $e(\hat{\mathcal{D}}^i_{val})$，其中 $e(\cdot)$ 表示所采用的主指标。结合 $\mathcal{D}_{val}$ 和 $\mathcal{A}_{val} = \{\hat{\mathcal{D}}^i_{val}\}_{i=1}^{T_2}$，我们可以计算任意两组之间的最大性能差异 α，

$$\alpha = \max\{\max|e(\mathcal{D}_{val}) - e(\hat{\mathcal{D}}^i_{val})|, \max|e(\hat{\mathcal{D}}^i_{val}) - e(\hat{\mathcal{D}}^j_{val}|)\}, \tag{2}$$

其中，$|\cdot|$表示绝对值运算。由于 $\mathcal{D}_{val}$ 和 $\mathcal{A}_{val}$ 是针对不同偏差环境模拟的，而理想的优化模型应在不同环境中保持稳定的性能，因此一个直观的想法是，α 值较小的模型更好。最后，根据 $e(\cdot)$是否为越高越好的指标，我们采用 $e(\mathcal{D}_{val}) - \alpha$ 或 $e(\mathcal{D}_{val}) + \alpha$ 作为修正后的性能结果ˆ，以便在风险可控的情况下，更好地离线选出最优模型 Θ^* 。

3.4 备注

基于IPS的方法是去偏推荐中的一个重要分支，但这类方法的性能高度依赖于逆倾向评分的估计精度。与之不同，我们的SSTE仅将逆倾向评分作为参考，用于模拟不同的偏差环境，因此对逆倾向评分的估计精度有更高的容忍度。我们的SSTE可应用于许多工业推荐场景，因为它不依赖于特定的数据、架构或训练方法，并且兼容大多数工业推荐模型。此外，由于采样得到的辅助子集通常远小于原始数据，即 $|\hat{\mathcal{D}}^i_{tr}| \ll m$ 和 $|\hat{\mathcal{D}}^i_{val}| \ll n$，我们的SSTE不会大幅增加时间和资源开销。

4 实验

本节首先介绍实验设置，然后开展广泛的实证研究，并展示我们提出的 SSTE 的有效性。

4.1 实验设置

数据集。 我们在实验中使用了一个非常常见的基准数据集和一个真实产品数据集，即 Yahoo! R3 [13] 和 Product。按照以往大多数关于 Yahoo! R3 的工作 [10,3],所采用的设置，我们首先将每个评分二值化，其中大于 3 的评分记为 $y = 1$，其余评分记为 $y = 0$。对于 Yahoo! R3 中的有偏反馈子集，我们按照 8∶2 的比例随机划分每个用户的反馈，得到训练集和验证集。Yahoo! R3 中的随机反馈子集全部用于无偏评估，因为它可以被视为由无偏场景生成的反馈。Product 是从腾讯理财通首页推荐业务所收集的日志数据中采样得到的一个子集，包含约 280 万用户、560 个物品和 960 万条反馈。根据反馈数据的不同属性，我们将其划分为有偏反馈子集和随机反馈子集。按照时间顺序对有偏反馈子集进行排序后

，我们以 8：2 的比例得到训练集和验证集。同样地，整个随机反馈子集均用于无偏评估。数据集的统计信息如表1所示。

表1. 数据集统计信息。P/N 表示正反馈与负反馈数量之比。

	Yahoo! R3		Product	
	#反馈	P/N (%)	#反馈	P/N (%)
训练集	254,713	67.02	7,133,519	5.21
验证集	56,991	67.00	1,633,716	5.20
测试集	54,000	9.64	870,158	4.34

评估指标。对于所有实验，我们采用推荐系统中广泛使用的四种评估指标，包括 ROC 曲线下面积（曲线下面积）、精确率（P@K）、召回率（R@K）和归一化折损累计增益（nDCG）。我们报告 K 为 5 和 10 时的 P@K 与 R@K 结果，以及 K 为 50 时的 nDCG 结果。由于曲线下面积是工业推荐中最常用的指标之一，我们选择它作为主要评估指标，用于为所有候选方法寻找最佳超参数。

基线方法。为了全面评估我们的 SSTE，在实验中，我们选取了一组仅基于有偏数据（即不使用随机化数据集）的代表性去偏推荐方法，包括逆倾向评分 [16], SNIPS [17], CVIB [18], AT [14], Rel [15] 和 DIB [11]。此外，与大多数先前研究类似，我们采用矩阵分解（MF） [6]和神经协同过滤（NCF） [4] 作为两种骨干模型。因此，这些基线方法中的每一种都有两个基于不同骨干模型的对应版本。请注意，我们未纳入 CausE [2], KDCRec [9] 和 AutoDebias [3], 因为它们需要随机化数据集来降低偏差。

实现细节。 所有候选方法均基于 TensorFlow 实现。我们将优化器设置为 Adam，并使用超参数搜索库 *Optuna* [1] ，以验证集上的曲线下面积为目标，加速超参数搜索过程。对于我们的 SSTE，考虑到性能与资源开销之间的权衡，我们将自采样的数量 T_1 和 T_2 均设置为 1。为避免在训练集上过拟合，我们还采用了耐心值为 5 的早停设置。我们分别在 $\{5, 10, \cdots, 195, 200\}$、$\left\{1e^{-5}, 1e^{-4}, \cdots, 1e^{-1}, 1\right\}$、$\left\{2^7, 2^8, \cdots, 2^{13}, 2^{14}\right\}$ 和 $\left\{1e^{-4}, 5e^{-4} \cdots 5e^{-2}, 1e^{-1}\right\}$ 的范围内调整嵌入维度、正则化权重、批大小和学习率。请注意，源代码可在 https://github.com/dgliu/DASFAA23_SSTE 获取。

4.2 总体结果

我们的 SSTE 与基线方法的比较结果如表2所示。当采用矩阵分解作为骨干模型时，如

表2上半部分所示，我们的 SSTE 在 Yahoo! R3 和 Product 两个数据集的所有指标上均始终优于所有基线方法。我们还可以观察到，大多数去偏方法都能在一定程度上提升推荐模型的无偏性能，其中 Rel 和 DIB 是最具竞争力的两种基线方法。与它们相比，我们的 SSTE 得益于自训练和自评估模块，能够取得更好的结果。当采用神经协同过滤作为骨干模型时，如表2下半部分所示，我们的 SSTE 在大多数情况下优于所有基线方法。我们发现，尽管在 Product 数据集上，我们的 SSTE 在 P@10 和 R@10 指标上略逊于 DIB，但在其他指标上仍具有明显优势，尤其是在主指标曲线下面积上。这可能是因为参数调优时仅考虑了曲线下面积。总体而言，我们的 SSTE 具有更好的无偏性能。

表2. 无偏评估的比较结果，其中最佳结果和次佳结果分别以粗体和下划线标示。曲线下面积是主要评估指标。

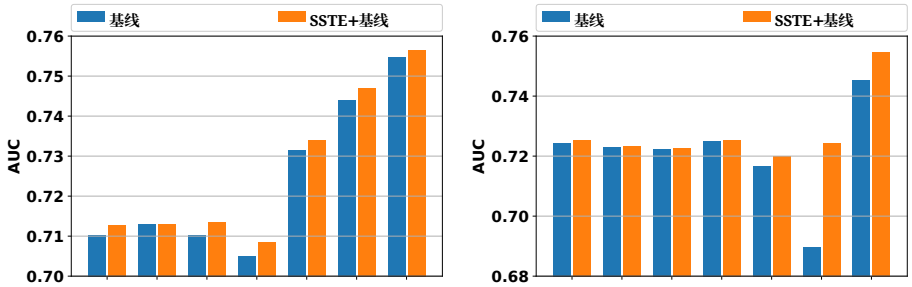
	Yahoo! R3						Product				
方法	曲线下面积	归一化折损累计增益	P@5	P@10	R@5	R@10	曲线下面积	归一化折损累计增益	P@5	P@10	R@5	R@10
MF	.7101	.0447	.0080	.0074	.0236	.0446	.8477	.0696	.0145	.0180	0687	.1707
IPS-MF	.7128	.0313	.0033	.0035	.0088	.0202	.8507	.0684	0139	.0180	.0663	.1717
SNIPS-MF	.7101	.0334	.0049	.0048	.0131	.0271	.8509	.0684	.0140	.0181	.0670	.1717
CVIB-MF	.7048	.0479	.0089	.0069	.0267	.0413	.8279	.0705	.0146	.0185	.0695	.1761
AT-MF	.7314	.0663	.0108	.0100	.0373	.0676	.8389	.0649	.0137	.0169	.0649	1595
Rel-MF	.7440	.0835	.0151	.0131	.0508	.0859	.8519	.0785	.0177	.0199	.0838	.1882
DIB-MF	.7547	.0920	.0169	.0145	.0538	.0930	.8510	.0781	.0159	.0202	.0764	.1922
SSTE-MF	**.7591**	**.0981**	**.0181**	**.0153**	**.0612**	**.0999**	**.8525**	**0878**	**.0222**	**.0211**	**.1062**	**.2010**

	Yahoo! R3						Product				
方法	曲线下面积	归一化折损累计增益	P@5	P@10	R@5	R@10	曲线下面积	归一化折损累计增益	P@5	P@10	R@5	R@10
NCF	.7244	.0294	.0022	.0026	.0068	.0147	.7930	.0795	.0192	.0195	.0907	.1845
逆倾向评分神经协同过滤	.7229	.0291	.0031	.0036	.0084	.0199	.8024	.0802	.0187	.0188	.0886	.1781
SNIPS-NCF	.7224	.0314	.0034	.0034	.0091	.0183	.8026	.0869	.0195	.0206	.0921	.1945
CVIB-NCF	.7250	.0393	.0053	.0044	.0150	.0250	.8034	.0776	.0170	.0193	.0805	.1831
AT-NCF	.7167	.0351	.0056	.0048	.0157	.0271	.7967	.0741	.0181	.0179	.0856	.1688
Rel-NCF	.6895	.0505	.0083	.0073	.0246	.0451	.7970	.0827	.0183	.0198	.0869	.1880
DIB-NCF	.7454	.0671	.0114	.0096	.0365	.0602	.8049	.1011	.0233	**.0241**	.1105	**.2293**
SSTE-NCF	**.7561**	**.0696**	**.0115**	**.0101**	**.0370**	**.0638**	**.8061**	**.1033**	**.0247**	.0228	**.1175**	.2171

4.3 兼容性分析

如第 3 节所述，由于我们的 SSTE 不依赖于特定的架构或训练方法，因此可以轻松地与现有的去偏推荐方法和传统推荐方法集成。为验证 SSTE 的兼容性，在实验中，我们将其与所有

基线方法集成，并在重新搜索最佳超参数后，将其与原始基线方法进行比较。我们在图 5 中报告了 Yahoo! R3 上的结果。可以发现，在大多数情况下，我们的 SSTE 都能显著提升不同基线方法的无偏性能。特别是，我们观察到，在基于逆倾向评分（IPS）的去偏推荐方法上，SSTE 的积极作用会减弱。这可能是因为对 IPS 的不准确估计会对推荐模型造成难以弥补的损害。



(a) 矩阵分解　　(b) 神经协同过滤

图 5. 我们的 SSTE 与不同基线方法在 Yahoo! R3 上的推荐结果。

4.4 在线 A/B测试

最后，我们将 SSTE 部署在腾讯理财通。腾讯理财通是一个大型互联网金融平台，致力于为用户提供优质的基金销售服务。该平台每天有数千万活跃用户，并产生和记录大量反馈日志。我们在首页推荐场景中进行了为期一个月的在线 A/B 测试。首页推荐场景是用户登录后的第一个页面。在这一推荐场景中，系统会向用户推荐一组基金，用户可以进行一些相关操作，例如跳过、点击和购买。展示页面如图 1 左侧所示。在线测试中的对比基准模型是经过精细调优的深度多任务模型，它分别预测点击量、转化量和购买金额。我们在相同的架构上部署 SSTE，并使用相同的训练数据集训练这两个模型。该数据集包含超过 3 亿条记录，涵盖两个月的反馈日志。在在线服务中，我们随机选择 10% 的用户作为实验组，由 SSTE 为其提供服务；另有 30% 的用户属于基准模型的对照组。与离线实验采用的评估指标不同，我们引入了金融推荐中更受关注的三项在线评估指标，即千次展示总点击量（CLPM）、千次展示总转化量（COPM）和千次展示购买金额（PAPM）。具体而言，CLPM、COPM 和 PAPM 分别可通过$\frac{Total\ Clicks}{Total\ Impressions} \times 1000$、$\frac{Total\ Conversions}{Total\ Impressions} \times 1000$ 和$\frac{Purchase\ Amount}{Total\ Impressions} \times 1000$计算得出，

$$\frac{}{} \qquad \frac{}{} \qquad \frac{}{}$$

在线 A/B 测试结果如图6所示。可以看出，我们的 SSTE 在三项评估指标上均带来了稳定提升。总体而言，在整个在线 A/B 测试期间，SSTE 在 CLPM、COPM 和 PAPM 上分别平均提升了 3.75%、7.20% 和 12.11%。这进一步证明了 SSTE 的有效性。

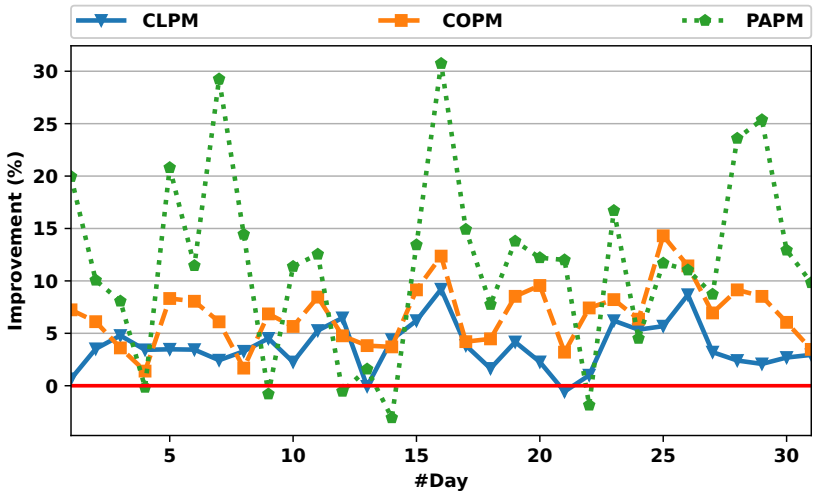


图6. 在在线 A/B 测试中，我们的 SSTE 相较于基准模型的提升，包括千次展示总点击数（CLPM）、千次展示总转化数（COPM）和千次展示购买金额（PAPM）。

5 结论

本文首先通过一项在线研究表明，在工业推荐应用中盲目去除所有偏差，未必总能带来预期的性能提升。为实现更好的准确率-偏差权衡，我们提出了一种简单而有效的自采样训练与评估（SSTE）框架，在去除有害偏差的同时保留有益偏差。SSTE 包含三个新模块：自采样模块构建用于训练和验证的去偏子集；自训练模块旨在基于原始训练数据和去偏子集联合学习准确率-偏差权衡；自评估模块旨在基于原始验证数据和去偏子集离线选出最优模型。我们在一个公开数据集和一个真实产品数据集上进行了大量实验，发现 SSTE 能有效提升推荐模型的无偏性能，且兼容性良好。最后，我们的 SSTE 在在线 A/B 测试中持续提升核心评估指标。

致谢。我们感谢中国国家自然科学基金（项目编号：61836005、62272315 和 62172283）的资助。

参考文献

1. Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M.：Optuna：新一代超参数优化框架。载于：SIGKDD，第2623–2631页（2019） 2. Bonner, S., Vasile, F.：用于推荐的因果嵌入。载于：RecSys，第104–112页（2018） 3. Chen, J., Dong, H., Qiu, Y., He, X., Xin, X., Chen, L., Lin, G., Yang, K.：AutoDebias：学习为推荐去偏。载于：SIGIR，第21–30页（2021） 4. He, X., Liao, L., Zhang, H., Nie, L., Hu, X., Chua, T.S.：神经协同过滤。载于：TheWebConf，第173–182页（2017） 5. Jadidinejad, A.H., Macdonald, C., Ounis, I.：推荐系统离线评估中的辛普森悖论。ACM TOIS 40(1)，1–22（2021） 6. Koren, Y., Bell, R., Volinsky, C.：推荐系统的矩阵分解技术。Computer (8)，30–37（2009） 7. Liang, D., Charlin, L., Blei, D.M.：推荐中的因果推断。载于：与第32届人工智能不确定性会议同期举办的“因果关系：从基础到应用”研讨会（2016） 8. Lim, D., McAuley, J., Lanckriet, G.：基于缺失隐式反馈的 Top-N 推荐。载于：RecSys，第309–312页（2015） 9. Liu, D., Cheng, P., Dong, Z., He, X., Pan, W., Ming, Z.：一种通过均匀数据进行反事实推荐的通用知识蒸馏框架。载于：SIGIR，第831–840页（2020） 10. Liu, D., Cheng, P., Zhu, H., Dong, Z., He, X., Pan, W., Ming, Z.：通过信息瓶颈缓解推荐中的混杂偏差。载于：RecSys，第351–360页（2021） 11. Liu, D., Cheng, P., Zhu, H., Dong, Z., He, X., Pan, W., Ming, Z.：通过信息瓶颈实现推荐中的去偏表征学习。ACM TORS（2022） 12. Liu, D., Lin, C., Zhang, Z., Xiao, Y., Tong, H.：推荐系统中的沉默的螺旋。载于：WSDM，第222–230页（2019） 13. Marlin, B.M., Zemel, R.S.：基于非随机缺失数据的协同预测与排序。载于：RecSys，第5–12页（2009） 14. Saito, Y.：针对非随机缺失显式反馈去偏的非对称三重训练。载于：SIGIR，第309–318页（2020） 15. Saito, Y., Yaginuma, S., Nishino, Y., Sakata, H., Nakata, K.：从非随机缺失隐式反馈中进行无偏推荐学习。载于：WSDM，第501–509页（2020） 16. Schnabel, T., Swaminathan, A., Singh, A., Chandak, N., Joachims, T.：将推荐视为干预：去偏学习与评估。载于：ICML，第1670–1679页（2016） 17. Swaminathan, A., Joachims, T.：反事实学习的自归一化估计量。载于：NeurIPS，第3231–3239页（2015） 18. Wang, Z., Chen, X., Wen, R., Huang, S.L., Kuruoglu, E.E., Zheng, Y.：从非随机缺失反馈中进行信息论反事实学习。载于：NeurIPS，第1854–1864页（2020） 19. Wang, Z., He, Y., Liu, J., Zou, W., Yu, P.S., Cui, P.：用于推荐通用去偏的不变偏好学习。载于：SIGKDD，第1969–1978页（2022） 20. Yang, L., Cui, Y., Xuan, Y., Wang, C., Belongie, S., Estrin, D.：针对非随机缺失隐式反馈的无偏离线推荐评估。载于：RecSys，第279–287页（2018）