

## 基于多任务学习的位置倾向性得分预测算法

曹泽麟<sup>1</sup> 徐君<sup>1,2</sup> 董振华<sup>3</sup> 文继荣<sup>1,2</sup>

<sup>1</sup>(中国人民大学高瓴人工智能学院 北京 100872)

<sup>2</sup>(大数据管理与分析方法研究北京市重点实验室(中国人民大学) 北京 100071)

<sup>3</sup>(华为技术有限公司诺亚方舟实验室 广东深圳 518129)

(zelincao@ruc.edu.cn)

## Prediction of the Positional Propensity Scores Based on Multi Task Learning

Cao Zelin<sup>1</sup>, Xu Jun<sup>1,2</sup>, Dong Zhenhua<sup>3</sup>, and Wen Jirong<sup>1,2</sup>

<sup>1</sup>(Gaoling School of Artificial Intelligence, Renmin University of China, Beijing 100872)

<sup>2</sup>(Beijing Key Laboratory of Big Data Management and Analysis Methods (Renmin University of China), Beijing 100071)

<sup>3</sup>(Huawei Noah's Ark Lab, Shenzhen, Guangdong 518129)

**Abstract** Users' click data distribution during search is quite different in different search scenarios. The existing methods such as CPBM (contextual position based model) only predict the positional propensity score in multiple scenarios through single model, which inevitably reduces the prediction accuracy in different scenarios and affects the effect of removing position bias. In this work, A MCPBM (multi-gate contextual position based model) based on multi-task learning is proposed. In this model, the information filtering structure is added to CPBM model to solve the problem of poor prediction accuracy during joint training on multi-scene data. At the same time, in order to alleviate the problem that the convergence speed of different tasks is inconsistent. We propose an exponentially weighted average dynamic adjustment algorithm, which speeds up MCPBM training and improves the overall prediction performance of MCPBM. The experimental results show that MCPBM model proposed in this paper is better than traditional CPBM model in prediction accuracy when multi-scene data is jointly trained. After using MCPBM model to remove the position bias in the training data, the ranking model obtained by training on the generated unbiased data promotes the *AvgRank* ranking metric of test data by 1%–5%.

**Key words** position bias; unbiased learning-to-rank; inverse probability weighting; multi-task learning; propensity score prediction

**摘要** 用户搜索时产生的点击数据分布,在不同的搜索场景下存在较大差异.现有算法如融合上下文的位置模型(contextual position based model, CPBM)往往只通过单个模型预测多种场景下的位置倾向性得分,不可避免地降低了模型在不同场景下的预测准确性,影响去除位置偏置的效果.基于上述问题提出一种基于多任务学习的多门专家混合位置倾向性得分预测模型(multi-gate contextual position based model, MCPBM),在CPBM模型的基础上加入信息筛选结构,解决了多场景数据联合训练时预测准确性不佳的问题.同时,为了缓解不同任务收敛速度不一致的问题,提出了指数加权平均权重动态调整算法,在加速模型训练的同时提升了模型整体预测性能.实验结果表明提出的MCPBM模型在多场景数据联合训练时,

收稿日期: 2021-08-20; 修回日期: 2022-02-11

基金项目: 国家重点研发计划项目(2019YFE0198200); 国家自然科学基金项目(61872338,61832017); 北京高校卓越青年科学家计划项目(BJJWZYJH012019100020098)

This work was supported by the National Key Research and Development Program of China (2019YFE0198200), the National Natural Science Foundation of China (61872338,61832017), and the Beijing Outstanding Young Scientists Program (BJJWZYJH012019100020098).

通信作者: 徐君(junxu@ruc.edu.cn)

预测准确性优于传统的 CPBM; 在使用 MCPBM 模型去除位置偏置后, 基于生成的无偏数据训练得到的排序模型, 在 AvgRank 排序指标上有 1%~5% 的提升.

**关键词** 位置偏置; 无偏排序; 逆概率加权; 多任务学习; 倾向性得分预测

**中图法分类号** TP311

在搜索场景下, 用户与系统的交互信息如点击日志将被记录, 并作为排序模型训练数据的重要来源. 用户点击日志具有时效性强、数据量大、反映用户真实喜好等优势, 同时这类数据中也存在大量的位置偏置<sup>[1]</sup>问题. 如何去除点击日志中的偏置信息, 使其能作为无偏排序模型的训练数据是提升模型排序指标和用户体验的关键方法之一. 目前业界主要是通过逆概率加权<sup>[2-4]</sup>算法去除位置偏置, 其关键在于获取各个位置准确的观测倾向性得分. 在实际应用场景中, 系统构建者往往通过融合上下文的位置模型 (contextual position based model, CPBM) 获取不同场景下各个位置观测倾向性得分<sup>[5]</sup>.

虽然 CPBM 模型已经取得了一定的去偏效果, 但在实际应用中不同搜索场景间的点击数据分布存在较大差异, 只通过一个 CPBM 模型预测多种场景位置倾向性得分的方式, 会不可避免地降低模型在各个场景的预测准确性. 例如, 用户会根据搜索目的提交导航查询<sup>[6]</sup> (navigational queries) 或者信息查询<sup>[6]</sup> (informational queries). 在导航查询中用户更倾向于浏览排序靠前的文档<sup>[7]</sup>, 因此其点击行为大多发生在序列头部; 而在信息查询中用户倾向于浏览更多的搜索结果, 用户对返回的搜索结果点击率明显高于导航查询, 点击位置也更加分散.

我们在真实的商业搜索引擎中观察到了上述情况, 图 1 中的统计数据来源于一个主流应用市场, 该应用市场每天服务千万级的用户. 从图 1 中可以看出, 在随机投放策略下, 用户在搜索精确词 (对应导航查询场景) 时各个位置的点击率与搜索泛词 (对应信息查询场景) 时有明显差异. 图 1 分析表明, 在不同搜索场景下用户的浏览行为和点击行为会有所不同. 使用混合多种场景分布的点击数据训练出的 CPBM 模型往往会受到数据之间分布不同的影响, 导致模型预测性能下降. 为每一种场景单独训练一个模型的方式, 虽然解决了上述问题, 但又会面临数据稀疏、无法利用数据之间共享信息的难题<sup>[8-9]</sup>.

针对 CPBM 模型受数据分布影响从而导致模型预测性能下降的问题, 本文提出一种多任务学习的多门专家混合位置倾向性得分预测模型 (multi-gate contextual position based model, MCPBM), 解决了单个

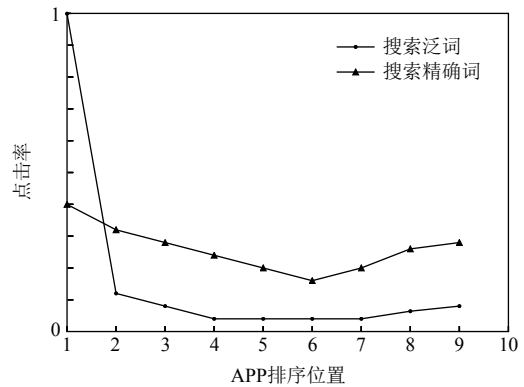


Fig. 1 Position click rate in random traffic

图 1 随机投放下的位置点击率

CPBM 模型在多场景数据联合训练时预测准确性不佳的问题. MCPBM 模型包含 1 个专家层 (expert layer) 网络和 1 个任务层 (task layer) 网络, 其中专家层网络用于提取场景的上下文特征, 任务层网络用于预测各个位置倾向性得分. MCPBM 在 2 层网络中分别设计了门共享结构进行信息筛选, 显式地对场景之间的相似性和差异性建模, 控制不同场景数据之间的信息共享程度, 从而综合利用不同场景中的共享信息帮助相似场景任务的学习.

本文工作的主要贡献包括 3 个方面:

1) 提出了基于多任务学习的多门专家混合位置倾向性得分预测模型 MCPBM, 解决多场景数据联合训练带来的模型预测性能下降问题.

2) 提出指数加权平均权重动态调整算法, 该算法通过对各个任务损失函数动态分配权重, 缓解不同任务收敛速度不一致的问题.

3) 实验结果表明, MCPBM 能够更加有效地去除点击日志数据中的位置偏置, 提高排序模型的排序指标; 实验分析表明, 本文提出的权重分配算法缓解了不同任务间收敛速度不一致的问题, 提升模型整体预测性能.

## 1 相关工作

### 1.1 位置偏置

位置偏置<sup>[1]</sup>是搜索场景中较为常见的一种偏置.

其定义为: 用户倾向与搜索结果列表中排序位置较高的文档进行交互, 并且用户的倾向性与文档是否满足用户的实际需求无关. 由于位置偏置的存在, 文档的点击率会随着展示位置的不同而不同, 产生排在较低位置文档的点击率明显低于排在较高位置文档的点击率的现象, 并且会导致训练数据中部分因排序位置较低缺少用户点击反馈的相关文档被当作负例进行训练. 同时位置偏置的存在会使模型高估排序位置较高文档的点击率, 低估排序位置较低文档的点击率, 使得位置较高的文档获得更多的反馈, 并在下一次展示中依然获得较为靠前的排序, 而排序位置较低没有获得用户反馈的相关文档在下一次的展示中依然被排在较低的位置. 因此为缓解上述现象, 需要对排序位置较低但是依然被点击的文档进行适当的权重调整, 使得模型可以捕捉到这部分代表性不足但与搜索内容相关的文档. 目前一个常用的解决方案是逆概率加权 (inverse probability weighting), 该算法去偏的关键在于获取各个位置准确的观测倾向性得分. 业界常通过随机数据 (uniform data)<sup>[10-11]</sup>、交换干预 (swap intervention)<sup>[12]</sup> 等数据干预方式来获取位置倾向性得分. 但现有干预算法在应用时会带来影响用户体验、降低产品业务指标、数据时效性不强等问题.

为解决上述问题, Agarwal 等人<sup>[4]</sup> 提出干预收割的数据干预方式, 利用了现实搜索场景中排序模型会被不断更新、各个模型之间存在一定差异性进行干预. 具体而言, 干预收割使用了多个排序模型处理相同搜索内容, 从而得到有差异搜索结果列表的数据集合. 此数据中存在同一个文档  $d$  既被排在位置  $k$  又被排在位置  $k'$  的情况,  $k \neq k'$ . 因此该算法关注用户与不同排序模型进行交互时文档排序的差异性, 来获取一个位置观测倾向性得分.

在位置模型<sup>[13]</sup> (position based model, PBM) 中, 文档  $d$  在位置  $k$  被观测到的概率仅由文档排序位置决定. 已有工作表明<sup>[5,14]</sup> 不同搜索场景下文档在各个位置的观测概率受文档排序位置、搜索内容上下文环境影响, 因此通过 PBM 模型获取的位置倾向性得分会与真实值之间存在一定偏差. 为提升 PBM 模型预测准确性, Fang 等人<sup>[5]</sup> 将用户搜索内容、文档特征等上下文信息加入 PBM, 得到的融合上下文的位置模型 CPBM 使其具备刻画不同场景下用户的点击行为, 从而提高了模型预测各个位置观测概率的准确性, 达到获取准确、无偏位置倾向性得分的目的. 本文对 CPBM 作了进一步的改进.

## 1.2 多任务学习

近年来, 多任务学习 (multi-task learning, MTL)<sup>[15-17]</sup> 已经在信息检索领域得到成功应用, 该学习算法通过捕捉各个任务之间共享信息和特有信息的方式来提高模型的泛化性能. Sheng 等人<sup>[9]</sup> 提出一种星式拓扑的多任务学习结构以及数据分区归一化的方式来进行多推荐场景数据联合学习, 取得了比单任务学习更好的效果. Zhao 等人<sup>[18]</sup> 将排序任务转换为多目标学习任务并在模型中加入辅助结构, 缓解训练数据中多种偏置带来的影响. Chen 等人<sup>[19]</sup> 提出一种梯度调整策略帮助模型获得更加稳定的收敛点. 本文通过多任务学习方式捕捉多场景点击数据中的共享信息, 进一步提升模型预测性能.

## 2 对 CPBM 的分析

CPBM 遵循了数据来自同一分布的假设, 然而现实中搜索系统收集的日志数据往往混合了多种用户行为分布的信息. 例如用户在搜索不同内容、浏览不同搜索结果列表时会展现出不同的观测行为和点击行为, 这些行为可以看作是来自不同用户行为分布的实例, 因此“数据来自同一分布”这一假设在现实场景中很难保证. 用包含多种分布的数据训练出的 CPBM 模型也会受到数据之间分布不同的影响, 面临模型预测性能下降等问题. 而为每一种分布数据训练一个 CPBM 模型的方式, 虽然消除了上述缺陷, 但又会面临训练数据稀疏的问题.

本文基于 Yahoo! Learning to Rank (简称 Yahoo) 数据集<sup>[20]</sup> 和 MQ2007 数据集<sup>[21]</sup> 对上述现象进行了初步的验证. 遵照 Fang 等人<sup>[5]</sup> 的实现方式, 本文构建了 4 个场景的用户点击数据, 每个场景数据特征由参数  $\theta$  刻画. 不同场景的数据之间存在一定相似性和差异性, 具体体现在参数  $\theta$  取不同值时各个位置观测概率的均值会随位置增大而减小, 但标准差会随参数  $\theta$  的增大而增大, 最终体现在不同参数  $\theta$  取值下各个位置具有不同的点击率.

表 1 展示了 CPBM 在 2 个数据集、3 个双场景设定下进行联合训练和独立训练时测试集错误率情况. 由于模型在  $\theta = 10$  场景计算出的错误率量级高于其他场景, 因此在计算后取对数 (lb). 可以看出, 在 Yahoo 数据集中, 在双场景 1 数据联合训练时模型在  $\theta = 0.1, \theta = 0.3$  的预测准确性相对于独立训练都有所提升; 在双场景 2 数据联合训练时模型在  $\theta = 0.1$  的预测准确性有所提升, 在  $\theta = 0.6$  的预测准确性有所

下降;在双场景3数据联合训练时模型在 $\theta = 0.1$ 和 $\theta = 10$ 的预测准确性都有所下降.在MQ2007数据集中,在双场景1数据联合训练时模型在 $\theta = 0.1$ 和 $\theta = 0.3$ 的预测准确性都有所提升;在双场景2数据和双场景3数据联合训练时模型在 $\theta = 0.1$ 的预测准确性有所提升, $\theta = 0.6$ 和 $\theta = 10$ 的预测准确性有所下降.以上实验结果表明了不同场景的数据之间存在提升模型预测能力的共享信息数据,也存在各个场景特有的信息数据,CPBM的预测性能会受联合训练时数据间分布不同的影响.因此为解决上述问题,需要在CPBM中加入具有信息筛选的结构.

**Table 1 Error Rate on the Test Set During Joint Training and Independent Training Under Dual Scene Data**

表1 双场景数据联合训练和独立训练时测试集错误率

数据集	训练方式	双场景1		双场景2		双场景3	
		$\theta = 0.1$	$\theta = 0.3$	$\theta = 0.1$	$\theta = 0.6$	$\theta = 0.1$	$\theta = 10$
Yahoo	联合	2.448	2.341	2.183	2.574	4.043	12.429
	独立	3.969	2.761	3.969	2.143	3.969	5.459
MQ2007	联合	1.824	2.660	1.892	4.434	1.964	34.714
	独立	2.226	3.223	2.226	4.431	2.226	32.302

### 3 MCPBM 分析

首先定义多场景日志数据.在多场景数据中, $T_i$ 代表其中1个场景的数据,在各个场景中 $\mathbf{X}_i$ 代表上下文特征向量, $\mathbf{Y}_i$ 代表真实标签,该标签可以是人为标注的相关性标签也可以是用户实际的点击标签. $T_i = \{\mathbf{X}_i, \mathbf{Y}_i\}$ . $\{T_i\}_{i=1}^n$ 代表所有场景数据的集合,其中有 $n$ 个场景的数据.

本文所提出的算法基于多任务学习的多门专家混合网络结构<sup>[22]</sup>(multi-gate mixture-of-experts, MMOE)进行构造,MMOE通过门共享结构显式地对任务之间的相似性和差异性建模,利用其他场景中的共享信息帮助相似场景任务的学习.MMOE将硬参数共享方式中共享底层结构<sup>[23]</sup>(share bottom)改进为多个专家层结构,避免了硬参数共享方式面临的负迁移、难以优化等问题.而相较于混合专家网络<sup>[24]</sup>(mixture-of-experts, MOE)中不同任务只使用1个门共享结构的方式来说,MMOE为每个任务都分配1个门共享结构,合理分配任务之间权重,提升了信息共享的灵活性.受MMOE结构的启发,本文提出多门专家混合位置倾向性得分预测模型MCPBM.在该模型中,1个场景的位置倾向性得分预测值由参与训练的所有场景

位置倾向性得分预测值加权得出,每个场景所占权重由场景独享的门共享结构给出.在该模型中,既利用各个场景特有的信息数据,也利用其他场景中的共享信息数据,解决CPBM受数据之间分布不同的影响导致预测准确性下降问题.

#### 3.1 专家层网络结构

图2展示了本文提出的MCPBM模型.在图2所示的网络结构中,下层为专家层网络,其中 $Expert_1, Expert_2, Expert_3$ 由多层感知机(multi-layer perceptron, MLP)网络构成,每个专家网络 $Expert_i$ 仅接收对应场景 $T_i$ 的上下文特征信息.从式(1)可以看出,专家层门共享结构用来捕捉特征层面的相似性,选择部分专家层的输出或者所有专家层的加权输出作为上层网络的输入.当其他场景数据与目标场景数据相关性越大时 $G_i$ 值越大,2种数据之间共享程度越高;当其他场景数据与目标场景数据相关性越小时 $G_i$ 值越小,2种数据之间共享程度越低.这种灵活的信息共享方式具有信息选择功能,可以将需要共享的信息传递到上层任务层网络中.1个专家网络输入对应 $T_i$ 场景的上下文信息 $\mathbf{X}_i$ ,输出为网络提取的特征信息 $\mathbf{X}'_i$ .1个专家层门共享结构输入为所有场景的上下文信息 $(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)$ ,输出为专家网络 $Expert_i$ 信息共享权重 $G_i$ 并且 $\sum_{i=1}^n G_i = 1$ .

$$G_i = \text{softmax}(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n), \quad (1)$$

因此专家层 $Expert_i$ 输出为

$$\mathbf{M}_i = \sum_{i=1}^n G_i \mathbf{X}'_i. \quad (2)$$

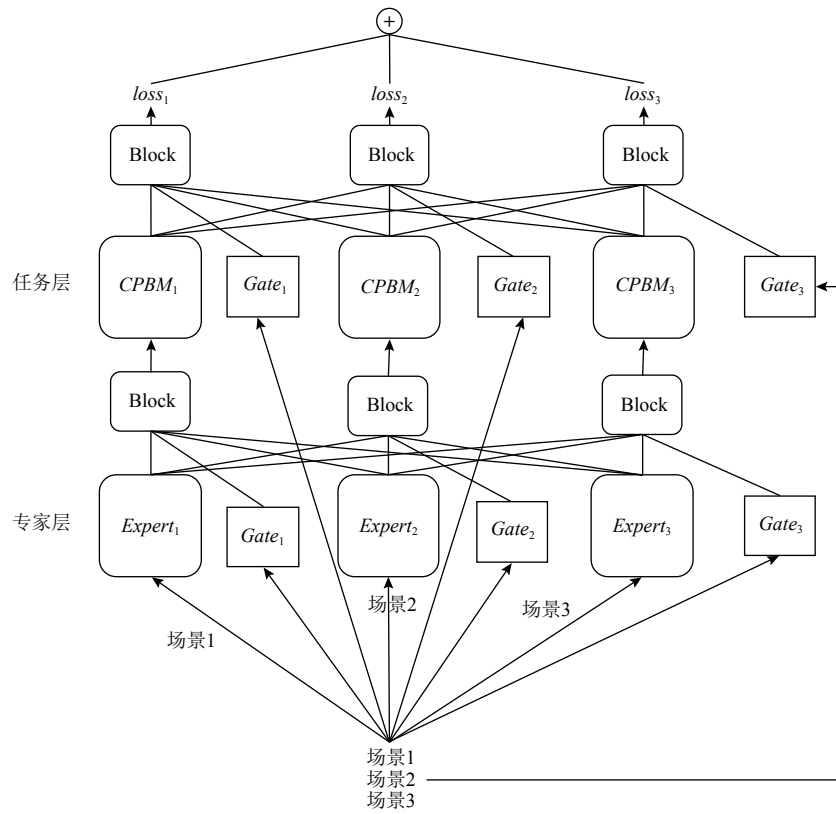
#### 3.2 任务层网络结构

上层为任务层网络,其中 $Task_1, Task_2, \dots, Task_n$ 由CPBM模型构成,用于获取每个位置的观测倾向性得分.1个专家网络与1个任务网络相对应,且2种网络的数量与场景个数 $n$ 相同. $Task_i$ 接收的特征信息由下层专家层和专家层门共享结构共同决定.任务网络的输出同样由任务层门共享结构和每个任务网络共同决定.每个任务网络的输出为 $\mathbf{H}_i$ ,1个任务层门共享结构输入为所有场景的上下文信息 $(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)$ ,输出为各个任务网络 $Task_i$ 信息共享权重 $Q_i$ 并且 $\sum_{i=1}^n Q_i = 1$ .

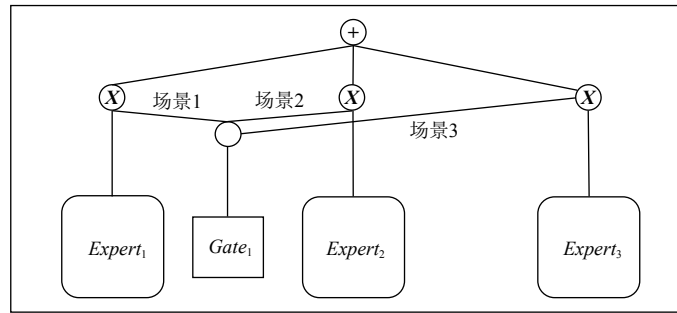
$$Q_i = \text{softmax}(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n), \quad (3)$$

因此任务层 $Task_i$ 输出为

$$\hat{\mathbf{Y}}_i = \sum_{i=1}^n Q_i \mathbf{H}_i. \quad (4)$$



(a) 多门专家混合位置倾向性得分预测模型



(b) Block信息融合模块内部结构

Fig. 2 Multi-gate contextual position based model and its structure of Block module

图2 多门专家混合位置倾向性得分预测模型及其Block模块结构

从式(3)可以看出,任务层门共享结构用来捕捉标签层面的相似性.值得注意的是,任务层与专家层中,门共享结构作用一致.在获取各个任务网络输出后,采用式(5)定义的损失函数:

$$loss_{total} = \sum_{i=1}^n \alpha_i loss_i, \quad (5)$$

其中,  $\alpha_i$  是各个损失函数的权重,采用3.3节的指数加权平均算法进行权重动态调整,  $loss_i$  采用式(6)定义的损失函数<sup>[5]</sup>:

$$loss = \arg \max_{p,r} \sum_{i \in I_{k,k'}} \sum_{k \neq k'} [C_{k,k'}^i \cdot \text{lb}(p(k, \mathbf{x}_i)r(k, k', \mathbf{x}_i)) + NC_{k,k'}^i \cdot \text{lb}(1 - p(k, \mathbf{x}_i)r(k, k', \mathbf{x}_i))], \quad (6)$$

式(6)采用了极大似然的估计方式.其中  $C_{k,k'}^i$  和  $NC_{k,k'}^i$  分别是  $I_{k,k'}$  干预组<sup>[5]</sup> 中文档加权点击率和加权未点击率,  $p(k, \mathbf{x}_i)$  和  $r(k, k', \mathbf{x}_i)$  分别是 CPBM 模型中的位置倾向性得分网络和相对相关性网络.

### 3.3 指数加权平均权重动态调整

由于不同场景各个位置点击率和位置倾向性得分有所不同,因此在使用不同场景数据训练时 MCPBM 有不同的收敛速度,存在某些场景损失函数收敛快而某些场景损失函数收敛慢的情况.而在 MCPBM 中,参与训练的所有数据决定模型最终的预测准确性.同时已有工作<sup>[25-27]</sup> 表明多任务学习模型性能依赖于每个任务损失函数之间的相对权重.为缓解不同场景任

务收敛速度不一致的问题,需要找到不同场景任务之间合适的组合方式,而通过网络搜索、人为调节等方式寻找最优权重组合十分不便.因此为解决上述问题,本文提出了指数加权平均权重动态调整法,具体加权方式见式(7)~(9):

$$\beta_i = \frac{\text{loss}_i(t-1)}{\text{loss}_i(t)}, \quad (7)$$

$$\sigma_i(t+1) = \frac{n \exp(\beta_i/S)}{\sum_{i=1}^n \exp(\beta_i/S)}, \quad (8)$$

$$\alpha_i(t+1) = \gamma \alpha_i(t) + (1-\gamma) \sigma_i(t+1). \quad (9)$$

式(7)中,  $\text{loss}_i(t)$ ,  $\text{loss}_i(t-1)$ 分别是任务  $i$  第  $t$  轮和第  $t-1$  轮的损失值,  $\beta_i$ 是任务  $i$  损失函数变化速率.式(8)中,  $S$ 是控制任务权重平滑度的超参,当  $S$ 取值越大时各个任务权重越接近.在获得  $\beta_i$ 后经过 softmax 运算并乘以任务数  $n$ 后得到  $t+1$  轮任务权重预测值  $\sigma_i(t+1)$ .为减少梯度下降和训练数据中的随机性,本文采用式(9)指数平均加权的方式对预测权重  $\alpha_i(t+1)$  进行加权,其中  $\gamma$ 是控制  $t$  轮之前任务权重在  $t+1$  轮中所占比例的超参,  $\gamma$ 值越大,任务  $t+1$  轮权重  $\alpha_i(t+1)$  与  $t$  轮之前任务权重越相关.

## 4 实验结果与分析

本节主要介绍实验中数据预处理方法、实验评估标准以及对实验结果进行分析,比较多任务学习模型和单任务学习模型在预测位置倾向性得分和位置去偏上的差异性,验证本文提出的 MCPBM 模型的有效性.

### 4.1 数据准备与预处理

本文采用 Yahoo 数据集和 MQ2007 数据集.由于 2 份公开数据集中不包含用户点击信息,因此需要对数据进行改造.本文遵照 Ai 等人<sup>[28]</sup>以及 Fang 等人<sup>[5]</sup>的实现方式生成模拟点击日志数据.值得注意的是,本文采用的用户点击数据模拟生成方式可以较好地刻画位置偏置,并且该方式在多个研究工作<sup>[2,5,28]</sup>中被广泛使用.同时,  $\theta$ 取值代表的场景一定程度上与我们在真实商业搜索引擎中收集的点击日志相符.

首先从训练集中随机抽取 1% 的搜索内容用于训练 2 个 SVM-Rank 排序模型<sup>[29]</sup>,模拟 1.1 节提到的干预收割数据干预方式中的多个排序算法,同时为了保证排序算法之间的相似性 1% 的训练数据中有 20% 搜索内容一致,其余的 80% 搜索内容不同.之后使用训练好的 2 个排序模型对剩余的训练数据进行

排序,得到文档排序位置,模拟用户搜索内容之后得到的结果列表.考虑到上下文信息对于观测概率的影响,本文使用数据集中的特征模拟现实场景中搜索内容的上下文信息.在该实验中,只关注前 10 个位置的偏置情况,因此每条搜索内容都会对应 1 个 10 维特征向量  $\mathbf{x}$ .向量  $\mathbf{x}$  由 2 部分组成:一部分  $(x^1, x^2, \dots, x^j)$  是从文档与搜索内容的特征中抽取;另一部分  $(x^{j+1}, x^{j+2}, \dots, x^{10})$  是从期望为 0、方差为 0.35 的正态分布生成.通过参数  $\delta = j/10$  来控制这 2 部分在 10 维特征向量中所占比重.

通过式(10)获取用户对于每个位置的观测概率并将生成的概率分布作为模型评估时的真实标签.

$$P(O = 1|k, \mathbf{x}) = \frac{1}{k^{\max(\mathbf{w}\mathbf{x}+1.0)}}, \quad (10)$$

其中,  $k$  是文档排序的位置,  $\mathbf{x}$  是抽取的上下文信息,  $\mathbf{w}$  是从  $[-\theta, \theta]$  的均匀分布中生成的向量.不同  $\theta$  取值代表不同场景数据.当  $\theta$  取较大值时,各个位置观测概率标准差较大,同时各个位置观测概率均值会随着排序位置的增大而减小.

在获取搜索内容上下文特征、各个位置的观测概率分布后,通过 CPBM 模型生成点击日志数据.本文采用 Fang 等人<sup>[5]</sup>使用的模拟点击方式并引入点击噪声来模拟现实场景中用户误点行为.用户点击相关文档的概率为 1,点击不相关文档的概率为 0.1.之后采用 1.1 节中介绍的干预收割数据干预方式获取干预数据.

### 4.2 实验评价指标

衡量模型预测各个位置观测倾向性得分的错误情况,采用式(11)的计算方式:

$$\text{error} = \frac{1}{|D|} \sum_{i \in D} \sum_{m=1}^k \text{abs} \left( \frac{\hat{p}_m(\mathbf{x}_i) - p_m(\mathbf{x}_i)}{p_m(\mathbf{x}_i)} \right), \quad (11)$$

其中  $k = 10$ , 该实验中只关注前 10 个位置的偏置情况,  $D$  是测试集,  $\hat{p}_m(\mathbf{x}_i)$  是  $m$  位置观测倾向性得分预测值,  $p_m(\mathbf{x}_i)$  是  $m$  位置观测倾向性得分真实值.模型预测位置观测倾向性得分越准确,计算出的  $\text{error}$  值越小.

在获得位置倾向性得分后,需要去除点击日志数据中的位置偏置并衡量使用去偏数据训练的排序模型的排序质量.因此在衡量排序质量时,首先使用训练好的位置倾向性得分预测模型对训练集数据进行去除位置偏置的处理,之后使用去偏数据训练基于倾向性得分的排序模型<sup>[2]</sup>(propensity SVM-Rank, PSR)并在测试集上进行评估,采用式(12)<sup>[2]</sup>来衡量排序模型的排序质量:

$$\text{AvgRank} = \frac{1}{|D^+|} \sum_{i \in D^+} \sum_{d \in Q(\mathbf{x}_i)} \text{rank}(d|f(\mathbf{x}_i)) \text{rel}(d, \mathbf{x}_i), \quad (12)$$

其中,  $f(x_i)$ 是训练得到的排序模型,  $rank(d|f(x_i))$ 是文档  $d$  排序的位置,  $rel(d, x_i)$ 是  $d$  与搜索内容  $x_i$  的相关度,  $\Omega(x_i)$ 是候选集文档集合.

### 4.3 实现细节

PyTorch 是一款开源的机器学习框架.本文所有实验使用 PyTorch 框架对模型进行实现.

#### 4.3.1 单任务学习模型

本文选取 CPBM、PBM、局部倾向性得分估计 (local estimators, LE)算法<sup>[4]</sup>、位置偏差感知学习框架 (position bias aware learning framework, PAL)<sup>[30]</sup>作为单任务学习的基线模型,该类模型使用单个  $\theta$  场景数据进行训练.其中 CPBM, PBM, LE 模型先通过预测位置倾向性得分再使用逆概率加权的方式去除偏置.CPBM 模型包括位置倾向性得分网络和相对相关性网络,这 2 种网络由 4 层 MLP 结构组成.表 2 展示了 CPBM 模型训练过程中选取的相关超参数.由于 PBM 模型未考虑上下文信息对于位置观测倾向性得分的影响,因此在获取干预数据后直接最大化式 (6).LE 通过干预数据中各个位置点击率来获取位置倾向性得分.PAL 框架采用位置偏置与点击率预测分开建模的方式,线上阶段仅使用去偏后的点击率预测模型进行预测.

Table 2 Hyperparameter Settings of Single-Task Learning

Model	
表 2 单任务学习模型的超参数设置	
超参数	取值
批次大小	{16,32,64}
学习率	[1E-4, 2E-4]
优化器	Adam
学习率递减	[50,100]

#### 4.3.2 多任务学习模型

多任务学习模型为 MCPBM 模型,该模型使用多个  $\theta$  场景数据进行联合训练.其中任务网络与单任务

Table 3 Hyperparameter Settings of Multi-Task Learning

Model	
表 3 多任务学习模型的超参数设置	
超参数	取值
批次大小	{16,32,64}
学习率	[1E-4, 3E-4]
优化器	Adam
任务权重比例 $\gamma$	[0.5, 0.7]
任务权重平滑度 $S$	[1, 3]
学习率递减	[30,50]

学习中 CPBM 模型网络结构、参数一致.专家网络由 4 层 MLP 结构组成,采用激活函数 ReLU;门共享结构由 2 层 MLP 结构组成并且最后一层采用 softmax 激活函数.表 3 展示了 MCPBM 训练过程中选取的相关超参数.

### 4.4 实验结果分析

本节主要是对 CPBM, PBM, LE, PAL, MCPBM 模型的预测位置倾向性得分以及去除位置偏置结果进行对比分析.

#### 4.4.1 位置倾向性得分预测结果分析

表 4 展示的实验结果是在  $\theta = 0.1, \theta = 0.3, \theta = 0.6$  以及  $\theta = 0.1, \theta = 0.6, \theta = 10$  的 3 场景数据下, MCPBM 模型和 3 种基线模型在测试集上预测位置观测倾向性得分的错误情况,采用式 (11) 的计算方式.由于模型在  $\theta = 10$  场景计算出的  $error$  量级高于其他场景,因此在计算后取对数 (lb).

Table 4 Error Rate on the Test Set Under Three Scene Data

表 4 3 场景数据下模型测试集错误率							
数据集	预测模型	3 场景 1			3 场景 2		
		$\theta = 0.1$	$\theta = 0.3$	$\theta = 0.6$	$\theta = 0.1$	$\theta = 0.6$	$\theta = 10$
Yahoo	LE	2.052	2.120	3.418	2.052	3.418	19.547
	PBM	0.458	1.419	2.909	0.458	2.909	20.220
	CPBM	3.969	2.761	2.143	3.969	2.143	5.459
	MCPBM	0.449	0.717	0.974	0.798	1.092	5.000
MQ2007	LE	1.503	1.574	2.955	1.503	2.955	34.543
	PBM	1.167	1.729	3.568	1.167	3.568	35.718
	CPBM	2.226	3.223	4.431	2.226	4.431	32.302
	MCPBM	1.125	1.485	2.872	1.991	2.431	29.767

实验结果表明, MCPBM 模型在各个场景的预测准确性都优于 3 种基线模型,仅在  $\theta = 0.1, \theta = 0.6, \theta = 10$  的 3 场景数据联合训练中,  $\theta = 0.1$  场景测试集错误率高于 PBM 模型.同时没有出现表 1 中 CPBM 模型在 2 种场景测试集上预测准确性都下降的情况.该实验结果表明了 MCPBM 模型具有一定的信息过滤能力,可以利用相似场景中的共享信息提升模型整体预测性能.

#### 4.4.2 指数加权平均权重动态调整结果分析

图 3~7 展示的实验结果是在 Yahoo 数据集中, MCPBM 模型采用指数加权平均权重动态调整算法、损失函数分配相等权重方式,以及本文参考的基线基于不确定性的权重调整算法 (uncertainty to weigh losses, Uncert)<sup>[27]</sup> 的情况下,分别在 4 场景数据和 3 场景数据联合训练时训练集的错误率变化曲线.由于 4

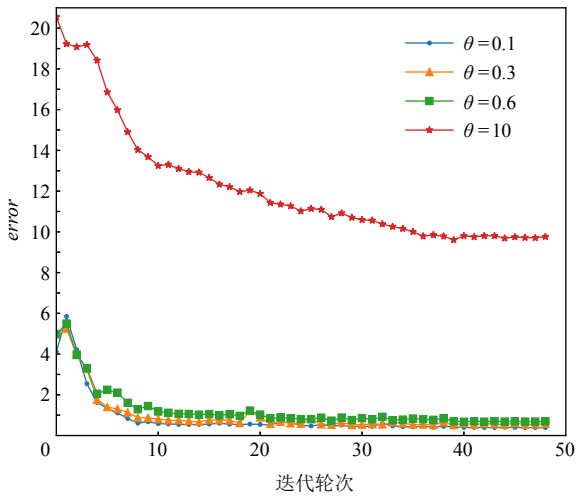


Fig. 3 Error rate curve of exponential weighted average in 4 scene data training

图3 指数加权平均在4场景数据训练时错误变化曲线

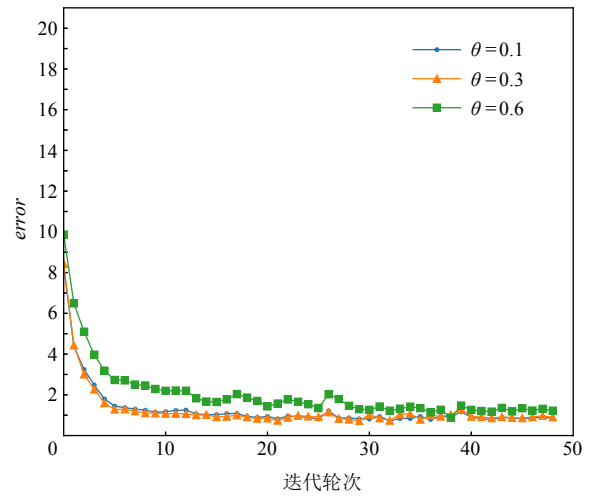


Fig. 6 Error rate curve of Uncert in 3 scene data training

图6 Uncert在3场景数据训练时错误变化曲线

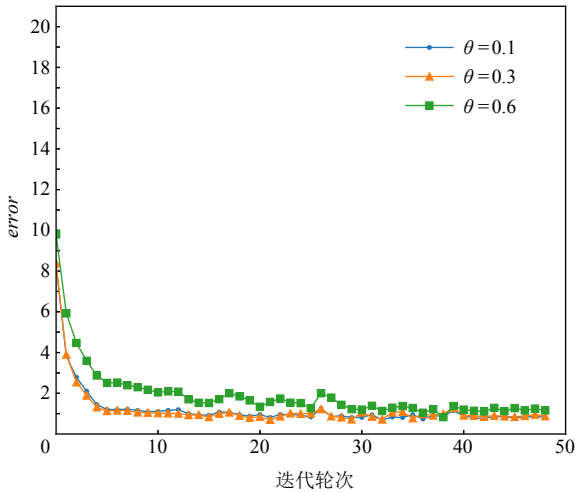


Fig. 4 Error rate curve of exponential weighted average in 3 scene data training

图4 指数加权平均在3场景数据训练时错误变化曲线

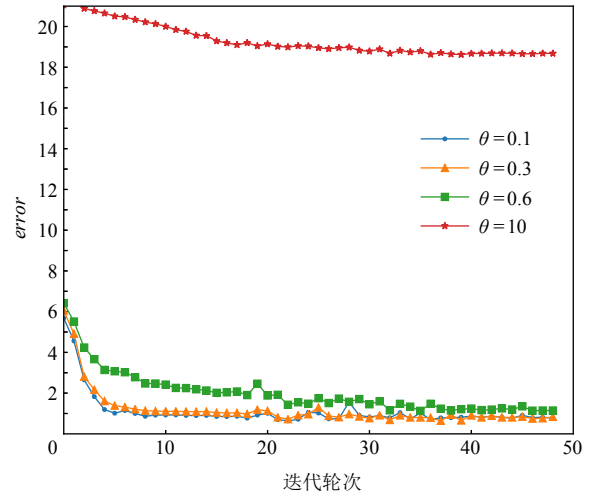


Fig. 7 Error rate curve of Uncert in 4 scene data training

图7 Uncert在4场景数据训练时错误变化曲线

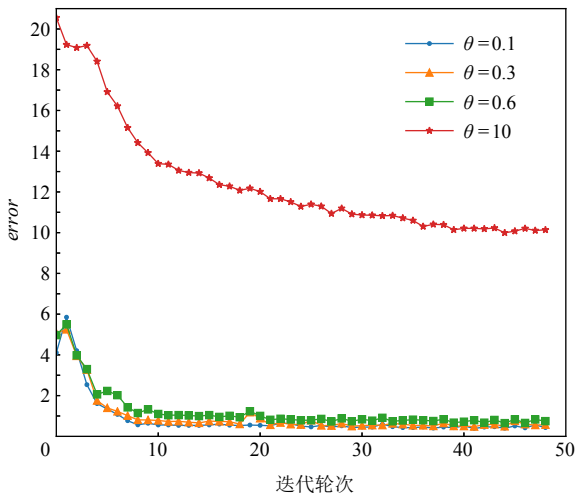


Fig. 5 Error rate curve of equal weight in 4 scene data training

图5 权重相等在4场景数据训练时错误变化曲线

场景数据联合训练时增加了1种场景的数据,因此模型在相同参数场景下错误率曲线会发生变化.例如,图3和图4中MCPBM在 $\theta = 0.1, \theta = 0.3, \theta = 0.6$ 的3场景训练集错误变化曲线有所不同.

实验结果表明本文提出的权重分配算法一定程度上缓解不同场景任务收敛速度不一致的问题,提升模型整体预测性能.在 $\theta = 0.1, \theta = 0.3, \theta = 0.6, \theta = 10$ 的4场景下使用指数加权平均权重动态调整的模型相较于分配相等权重的模型在训练集的error值上有4.1%的下降.同时文本提出的权重分配算法明显优于基线Uncert的权重调整算法.

#### 4.4.3 模型去偏结果分析

为验证PBM, CPBM, LE, PAL, MCPBM模型去除位置偏置的效果,本文使用PBM, CPBM, LE, MCPBM模型去偏后的数据和Click未去偏点击数据训练PSR



模型并在测试集上采用 *AvgRank* 指标评估排序模型的排序质量。

从表 5 可以看出, MCPBM 去偏效果优于 4 种位置倾向性得分预测去偏模型, 在 *AvgRank* 指标上均有 1%~5% 的提升. 实验结果验证了 MCPBM 模型能较较好地去除点击日志数据中的位置偏置问题. 从直接使用 Click 数据训练得到排序模型的排序指标来看, 位置偏置会严重影响排序模型的排序质量. 同时表 4 的实验结果表明在使用逆概率加权算法时获得一个准确的位置倾向性得分是去除位置偏置的关键所在.

**Table 5 AvgRank on the Test Set Under Three Scene Data**

表 5 3 场景数据下测试集的 *AvgRank* 情况

数据集	预测模型	3 场景		
		$\theta = 0.1$	$\theta = 0.3$	$\theta = 0.6$
Yahoo	LE	18.34	18.28	18.44
	PBM	18.27	18.25	18.43
	CPBM	18.42	18.32	18.39
	PAL	18.30	18.29	18.33
	MCPBM	18.26	18.25	18.36
	Click	18.51	18.47	18.57
MQ2007	LE	16.86	16.84	16.79
	PBM	16.84	16.86	16.80
	CPBM	16.88	16.91	16.82
	PAL	18.28	17.59	17.83
	MCPBM	16.82	16.83	16.74
	Click	16.89	17.16	16.96

## 5 总结与展望

本文提出一种基于多任务学习的位置倾向性得分预测模型和指数加权平均权重动态调整算法, 解决了已有的 CPBM 模型在多场景数据联合训练时, 因数据之间分布不同而导致模型预测性能下降的问题, 从而更准确地估计了不同场景下位置观测倾向性得分. 实验结果表明, 与 CPBM 相比, MCPBM 模型在多场景搜索中具有更优的去偏效果, 有效缓解了点击日志中的位置偏置, 提升了排序模型的排序质量.

在下一步工作中, 我们将尝试改进更多去除位置偏置的算法以及设计去偏任务和排序任务联合训练的框架.

致谢 感谢中国人民大学公共政策实验室的支持.

**作者贡献声明** 曹泽麟负责所有实验、数据分析, 以及文章的撰写; 徐君对本文选题、组织结构和文章写作提供了关键性的指导意见; 董振华对本文组织结构和部分内容提供了重要的指导意见; 文继荣对本文的选题提供了重要的指导意见.

## 参 考 文 献

- [1] Chen Jiawei, Dong Hande, Wang Xiang, et al. Bias and debias in recommender system: A survey and future directions [J]. arXiv preprint, arXiv: 2010.03240, 2020
- [2] Joachims T, Swaminathan A, Schnabel T. Unbiased learning-to-rank with biased feedback[C] //Proc of the 10th ACM Int Conf on Web Search and Data Mining. New York: ACM, 2017: 781-789
- [3] Hu Ziniu, Wang Yang, Peng Qu, et al. Unbiased lambdamart: An unbiased pairwise learning-to-rank algorithm[C] //Proc of the 28th World Wide Web Conf. New York: ACM, 2019: 2830-2836
- [4] Agarwal A, Zaitsev I, Wang Xuanhui, et al. Estimating position bias without intrusive interventions[C] //Proc of the 12th ACM Int Conf on Web Search and Data Mining. New York: ACM, 2019: 474-482
- [5] Fang Zhichong, Agarwal A, Joachims T. Intervention harvesting for context-dependent examination-bias estimation[C] //Proc of the 42nd Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2019: 825-834
- [6] Broder A. A taxonomy of web search[J]. ACM Special Interest Group on Information Retrieval Forum, 2002, 36(2): 3-10
- [7] Dupret G E, Piwowarski B. A user browsing model to predict search engine click data from past observations[C] //Proc of the 31st Annual Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2008: 331-338
- [8] Li Pengcheng, Li Runze, Da Qing, et al. Improving multi-scenario learning to rank in e-commerce by exploiting task relationships in the label space[C] //Proc of the 29th ACM Int Conf on Information & Knowledge Management. New York: ACM, 2020: 2605-2612
- [9] Sheng Xiangrong, Zhao Liqin, Zhou Guorui, et al. One model to serve all: Star topology adaptive recommender for multi-domain CTR prediction [J]. arXiv preprint, arXiv: 2101.11427, 2021
- [10] Yuan Bowen, Hsia J Y, Yang Mengyuan, et al. Improving ad click prediction by considering non-displayed events[C] //Proc of the 28th ACM Int Conf on Information and Knowledge Management. New York: ACM, 2019: 329-338
- [11] Yuan Bowen, Liu Yaxu, Hsia J Y, et al. Unbiased ad click prediction for position-aware advertising systems[C] //Proc of the 14th ACM Conf on Recommender Systems. New York: ACM, 2020: 368-377
- [12] Wang Xuanhui, Golbandi N, Bendersky M, et al. Position bias estimation for unbiased learning to rank in personal search[C] //Proc of the 11th ACM Int Conf on Web Search and Data Mining. New York: ACM, 2018: 610-618
- [13] Chuklin A, Markov I, Rijke M. Click models for web search[J]. Synthesis Lectures on Information Concepts, Retrieval, and Services,

- 2015, 7(3): 1–115
- [14] Chandar P, Carterette B. Estimating clickthrough bias in the cascade model[C] //Proc of the 27th ACM Int Conf on Information and Knowledge Management. New York: ACM, 2018: 1587–1590
- [15] Thung K H, Wee C Y. A brief review on multi-task learning[J]. *Multimedia Tools and Applications*, 2018, 77(22): 29705–29725
- [16] Vandenhende S, Georgoulis S, Van Gansbeke W, et al. Multi-task learning for dense prediction tasks: A survey[J/OL]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021 [2021-12-09]. <https://arxiv.org/pdf/2004.13379.pdf>
- [17] Wu Bin, Lou Zhengzheng, Ye Yangdong. A collaborative filtering recommendation algorithm for multi-source heterogeneous data[J]. *Journal of Computer Research and Development*, 2019, 56(5): 1034–1047 (in Chinese)  
(吴宾, 娄铮铮, 叶阳东. 一种面向多源异构数据的协同过滤推荐算法[J]. *计算机研究与发展*, 2019, 56(5): 1034–1047)
- [18] Zhao Zhe, Hong Lichan, Wei Li, et al. Recommending what video to watch next: A multitask ranking system[C] //Proc of the 13th ACM Conf on Recommender Systems. New York: ACM, 2019: 43–51
- [19] Chen Zhao, Ngiam J, Huang Yanping, et al. Just pick a sign: Optimizing deep multitask models with gradient sign dropout [J]. arXiv preprint, arXiv: 2010.06808, 2020
- [20] Chapelle O, Chang Yi. Yahoo! Learning to rank challenge overview [C/OL] //Proc of the 2010 Int Conf on Yahoo! Learning to Rank Challenge. Cambridge, MA: JMLR, [2011-12-09]. <http://proceedings.mlr.press/v14/chapelle11a>
- [21] Qin Tao, Liu Tiejian. Introducing LETOR 4.0 datasets [J]. arXiv preprint, arXiv: 1306.2597, 2013
- [22] Ma Jiaqi, Zhao Zhe, Yi Xinyang, et al. Modeling task relationships in multi-task learning with multi-gate mixture-of-experts[C] //Proc of the 24th ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining. New York: ACM, 2018: 1930–1939
- [23] Caruana R. Multitask learning[J]. *Machine Learning*, 1997, 28(1): 41–75
- [24] Jacobs R A, Jordan M I, Nowlan S J, et al. Adaptive mixtures of local experts[J]. *Neural Computation*, 1991, 3(1): 79–87
- [25] Chen Zhao, Badrinarayanan V, Lee C Y, et al. Gradnorm: Gradient normalization for adaptive loss balancing in deep multitask networks[C] //Proc of the 35th Int Conf on Machine Learning. Cambridge, MA: JMLR, 2018: 794–803
- [26] Liu Shikun, Johns E, Davison A J. End-to-end multi-task learning with attention[C] //Proc of the 32nd of the IEEE/CVF Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2019: 1871–1880
- [27] Kendall A, Gal Y, Cipolla R. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics[C] //Proc of the 31st IEEE Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2018: 7482–7491
- [28] Ai Qingyao, Bi Keping, Luo Cheng, et al. Unbiased learning to rank with unbiased propensity estimation[C] //Proc of the 41st Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2018: 385–394
- [29] Joachims T. Optimizing search engines using clickthrough data[C] //Proc of the 8th ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2002: 133–142
- [30] Guo Huifeng, Yu Jinkai, Liu Qing, et al. PAL: A position-bias aware learning framework for CTR prediction in live recommender systems[C] //Proc of the 13th ACM Conf on Recommender Systems. New York: ACM, 2019: 452–456



**Cao Zelin**, born in 1998. PhD candidate. His main research interests include data mining and recommendation system.

曹泽麟, 1998年生.博士研究生.主要研究方向为数据挖掘和推荐系统.



**Xu Jun**, born in 1979. PhD, professor, PhD supervisor. His main research interests include information retrieval, Internet search, machine learning and big data analysis.

徐君, 1979年生.博士,教授,博士生导师.主要研究方向为信息检索、互联网搜索、机器学习 and 大数据分析.



**Dong Zhenhua**, born in 1984. PhD, professor. His main research interests include recommendation system, counterfactual learning and mobile computing.

董振华, 1984年生.博士,研究员.主要研究方向为推荐系统、反事实学习和移动计算.



**Wen Jirong**, born in 1972. PhD, professor, PhD supervisor. His main research interests include Internet big data management, information retrieval, data mining and machine learning.

文继荣, 1972年生.博士,教授,博士生导师.主要研究方向为互联网大数据管理、信息检索、数据挖掘和机器学习.