

阿里2023-创新点击率预测，深度意图感知网络掀起新浪潮



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

19 人赞同了该文章

Introduction

小程序⁺是电商平台的新营销工具，如Juhuasuan、Little Black Box和Shoppee的闪购等，每日贡献数十亿页面浏览量⁺。用户通过点击App内的入口图标进入小程序，如图(a)所示。点击后，用户会看到一些推荐项，如图(b)所示。

本研究关注的是在小程序中推荐后续项的任务，即触发诱导推荐在min-app中的TIRA。传统的点击率⁺预测模型如DIN、DIEN、SIM和GIFT无法有效处理触发项效果。min-app的用户通常对特定min-app有强烈的粘性，他们的行为更多地受到日常购物习惯的影响，而不是触发的影响。因此，过度强调触发重要性的推荐方法可能导致推荐结果过于依赖触发项，不利于min-app的常规用户。



Figure 1: Trigger-Induced Recommendation in Mini-Apps

对于常规用户，推荐结果需要具备多样性，以满足他们在min-app上的多种兴趣。

1.提出了解决TIRA问题的深度意向感知网络 (DIAN) 模型，通过引入auxiliary task和posterior labeling策略提高了用户进入意图的准确性。

2.DIAN模型由意图网、有意识触发感知网和无意识触发感知网三部分组成。意图网用于识别用户进入意图，有意识触发感知网和无意识触发感知网分别评估在触发条件下的CTR。

3.DIAN在大规模真实世界的实验数据上表现出色，为Juhuasuan这款知名淘宝小程序带来了在线Item Page View量提升9.39%和点击率提升4.74%的收益，目前已经部署在淘宝App，服务于Juhuasuan的所有流量。

TIRA Problem Formulation

1. 用户进入意图的估计：利用 $Intent = 0$ 和 $Intent = 1$ 表示用户是否进入min-app或trigger-item，进而可以表示用户进入意图的概率分布⁺

$$(P(Intent = 0), P(Intent = 1))$$

因为只有两种原因进入min-app，所以 $P(Intent = 0) + P(Intent = 1) = 1$ 。

2. 基于用户进入意图到trigger-item的CTR估计：独立估计用户进入意图和基于用户进入意图到trigger-item的CTR。

3. 基于用户进入意图到min-app的CTR估计：独立估计用户进入意图和基于用户进入意图到min-app的CTR。通过这种方式，CTR预测任务可以转换为另外一种形式进行估计，即分别估计用户进入意图、基于用户进入意图到trigger-item的CTR以及基于用户进入意图到min-app的CTR。

$$P(Click | Target) = P(Click | Target, Intent = 0) * P(Intent = 0) + P(Click | Target, Intent = 1) * P(Intent = 1)$$

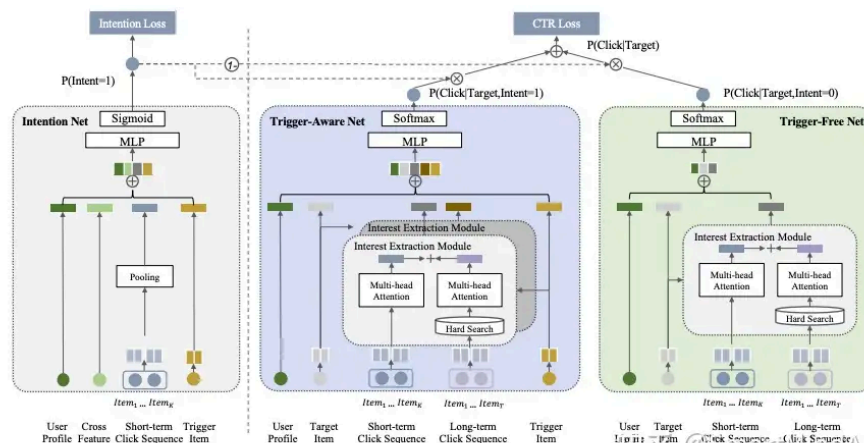


Figure 3: DIAN consists of three sub-nets: Intention Net, Trigger-Aware Net and Trigger-Free Net.

Deep Intention-Aware Network 我们设计针对给定两个不同意图时用户对目标项的兴趣的单独神经网络⁺，并利用估计的用户进入意图分数来平衡两种不同类型用户兴趣，提出了DIAN网络，(DIAN)是一种用于处理TIRA问题的深度学习模型，它采用了三个关键子网：IntentionNet、Trigger-Aware Net和Trigger-Free Net。

1. **IntentionNet**：此子网的目标是估计用户进入意图，即用户是否会进入min-app。其工作原理是将用户行为轨迹和触发项作为输入，通过多层神经网络⁺学习复杂的函数关系，进而构建用户进入意图网络。该子网使用的激活函数⁺通常为ReLU或者Sigmoid，以增加模型的稳定性和收敛速度⁺。通过训练，Intention Net可以很好地估计用户进入意图的概率分布。

2. **Trigger-Aware Net**：此子网的目标是估计基于用户进入意图到trigger-item的CTR，即在用户进入min-app后，点击触发项的概率。为此，需要使用到观测器，即对触发项是否被用户点击进行标记。通过计算这些标记的概率分布，即可得到基于用户进入意图到trigger-item的CTR。该子网同样采用多层神经网络，并且需要使用ReLU或者Sigmoid激活函数。

3. **Trigger-Free Net**：此子网的目标是估计基于用户进入意图到min-app的CTR，即在用户进入min-app后，直接访问min-app的概率。为此，需要对触发项的相关属性进行编码，并将其作为一个额外的输入添加到触发基础的模型中。通过训练，Trigger-Free Net可以估算出基于用户进入意图到min-app的CTR。

Intention Net

$$\hat{y}_{Intent} = \text{Sigmoid}(\text{MLP}(H_{user} \oplus H_{seq} \oplus H_{crs} \oplus H_{tri}))$$

Trigger-Aware Net

1. **输入部分**：User profile特征，User短序列点击项，User长序列点击项，Trigger项特征，Target项特征。同样地，我们在Intention Net中采用了嵌入技术将这些原始特征投影到嵌入向量中，并将它们分别表示为 E_{user} ， E_{sseq} ， E_{lseq} ， E_{tri} 和 E_{tar} 。

$$E_{sseq} = [E_{item}^{s1}, \dots, E_{item}^{sK}]$$

$$E_{lseq} = [E_{item}^{l1}, \dots, E_{item}^{lT}]$$

2. **双层兴趣提取**：根据Trigger项和Target项生成用户兴趣表示 h_{tri} 和 h_{tar} 。对于Trigger项和Target项，我们分别产生用户兴趣表示 h_{tri} 和 h_{tar} ，以防止它们在不同的情况下相互干扰。

$$H_{tar} = \text{InterestExtract}_{tar}(E_{tar}, E_{sseq}, E_{lseq})$$

举例来说，我们引入了一个层的Interest Extraction (IE)，该层由短时IE和长时IE模块组成，其中短时IE模块处理过去14天的点击序列，长时IE处理过去180天的点击序列。对于短时IE，我们使用多头目标注意力与Trigger项相关的短时点击序列中检索相关信息，从而生成用户短时兴趣表示 H_{tri}^s ：

$$H_{tri}^s = \text{MultiHeadAtt}(E_{sseq}, E_{tri})$$

其中 H_{tri}^s 是用户短时兴趣表示， E_{sseq} 是用户短时点击序列的嵌入向量， E_{tri} 是Trigger项的嵌入向量⁺。要注意的是，由于用户的兴趣在短时间内可能有所变化，因此我们只使用过去14天的点击序列来计算用户短时兴趣。

$$head_i = W_i^V E_{sseq} \cdot \text{Softmax}(W_i^Q E_{tri} \odot W_i^K \cdot E_{sseq})$$

$$H_{tri}^s = [head_1 \oplus head_2 \dots \oplus head_n]$$

$$H_{tri}^l = \text{MultiHeadAtt}(E_{lseq}, S_{lseq})$$

其中， S_{lseq} 是最近 K 个属于触发项叶子类别相同的长时点击项， H_{tri}^l 是针对触发项的长时兴趣表示， E_{lseq} 是长时点击序列的嵌入向量， S_{lseq} 是过滤后的长时点击序列。和短时IE一样，我们也使用一个软查询的目标注意力函数来从过滤后的长时点击序列中柔和地检索相关信息，从而生成长时兴趣表示：

$$H_{tri}^l = \text{SoftQueryAtt}(E_{lseq}, S_{lseq}, \lambda)$$

其中，SoftQueryAtt是SIM提出的搜索型目标注意力函数，它将长时点击序列划分为多个子集并计算每部分的相关性权重，然后选取权重最大的子集作为结果，可以有效降低长时点击序列的长度。需要注意的是，当构建长时兴趣表示时，需要使用 K 个最近的属于同一叶子类别的长时点击项。

$$head_j = W_j^V S_{lseq} \cdot \text{Softmax}(W_j^Q E_{tri} \odot W_j^K \cdot S_{lseq})$$

$$H_{tri}^l = [head_1 \oplus head_2 \dots \oplus head_n]$$

需要注意的是，长短期用户兴趣表示的投影矩阵⁺不共享，为了简化，我们可以省略下标 $sseq$ 和 $lseq$ 。通过添加短时和长时用户兴趣表示，我们可以得到用户对触发项的兴趣表示，即

$$H_{tri} = H_{tri}^{sseq} + H_{tri}^{lseq}$$

对于目标项，我们按照同样的方式进行计算，可以得到用户对目标项的兴趣表示 H_{tar} 。最后，我们将 E_{user} ， E_{tri} ， E_{tar} ， H_{tri} 和 H_{tar} 组合在一起并通过全连接层⁺进行处理：

$P(\text{Click} | \text{Target}, \text{Intent} = 1)$ 。其中， \hat{y}_{tan} 表示点击目标物品的CTR。

Trigger-Free Net

图展示了TFN的基本架构，其基本原理是在不考虑触发项的情况下估计点击率。我们采用类似TAN的方法，但是将触发项的特征从输入中删除，同时移除了兴趣提取层，使其退化为SIM模型。为了减少参数数量，我们强制使TFN与TAN共享相同的嵌入层，因为嵌入层占CTR模型参数的大部分。

同样地，我们按照公式的方式生成用户对目标项的兴趣表示 H_{tar}' 。我们将三个嵌入向量组合在一起并通过全连接层进行处理：在TFN模型中，我们通过模型预测用户的点击行为，给定用户进入意图是与迷你应用相关联的，即 $P(\text{Click} | \text{Target}, \text{Intent} = 0)$ 。其中， \hat{y}_{tfn} 表示点击目标物品的CTR。

根据估算的用户意图 \hat{y}_{int} 以及给定不同意图下的点击率 \hat{y}_{tan} 和 \hat{y}_{tfn} ，我们可以根据Section中的公式 (Eq.) 来估计最终的点击率。

$$\hat{y} = \hat{y}_{int} * \hat{y}_{tan} + (1 - \hat{y}_{int}) * \hat{y}_{tfn}$$

最终的**目标函数**[†]由点击率损失和意图损失组成：

CTR损失: $\ell_{ctr} = \frac{1}{n} \sum_{i=1}^n -\log \hat{y}_i$

目标意图损失: $\ell_{intent} = \frac{1}{m} \sum_{j=1}^m (y_j - \hat{y}_j)^2$

$$L = -\frac{1}{|\mathcal{D}|} \sum_{(\hat{y}, y) \in \mathcal{D}} \left(y \log \hat{y} + (1 - y) \log(1 - \hat{y}) \right) + \alpha (y_{int} \log \hat{y}_{int} + (1 - y_{int}) \log(1 - \hat{y}_{int}))$$

其中， \mathcal{D} 是训练集， $y \in \{0, 1\}$ 是点击率预测任务的标签， $y_{int} \in \{0, 1\}$ 是用户意图估计算法的任务标签，即表示用户是否以迷你应用或触发项作为进入意图。 α 是一个**超参数**[†]，用于平衡点击率损失和意图损失，我们在实验中将其设置为0.1。

Experiments

Experimental Setup

数据集：本文使用了淘宝Juhuasuan上的一个著名的小程序来收集数据集，该小程序主要销售品牌折扣产品。首先，从2022年8月1日到2022年9月1日，从用户的印象和点击日志中抽取一部分样本作为训练集，然后在2022年9月2日抽取**测试集**[†]，这样可以严格模拟在线服务环境。训练集中有10430212219个样本，测试集中有352959879个样本，总共涉及173406246个用户和6500833个商品。

基准方法：本文比较DIAN与其他两种行业常用的基准方法，一种是无触发项方法，包括Wide & Deep、DIN、SIM；另一种是有触发项的方法，包括DIHN和我们的Trigger-Aware Net (TAN)。

实现方式：所有竞争对手的隐藏层维度都是1024×512×256。对于序列模型，只使用短期点击序列进行DIN，而同时使用短期和长期序列进行SIM (TFN)、DIHN、TAN和DIAN。Adam优化器被用于学习率为0.01。对于所有注意力模块，注意力头的数量都被设置为4。

Table 1: CTR prediction comparison on the testing set

Method	AUC
Wide & Deep	0.767
DIN	0.791
SIM (TFN)	0.800
DIHN	0.805
TAN	0.806
DIAN	0.812
DIAN (remove intention loss)	0.808
DIAN (remove Intention Net)	0.806

Online A/B Testing

实施部署：我们实现了并部署了提议的DIAN Mini-app。从2022年9月16日至10月15日，在淘宝App上进行了严格的在线A/B测试，比较DIAN和以前的在线模型SIM。表展示了比较结果（按浏览量 (PV)、单个页面浏览量 (IPV) 和点击率 (CTR) 进行比较），其中PV和IPV是用户在进入迷你应用程序后浏览和点击的平均数量。

Method	IPV	PV	CTR
SIM	0.555	23.93	2.32%
DIAN	0.607	24.91	2.43%
Lift rate	+9.39%	+4.10%	+4.74%

明显可以看出，DIAN比SIM有着显著的优势：DIAN将贡献高达4.74%的CTR提升、4.10%的PV提升和9.39%的IPV提升，这表明用户不仅会浏览更多，而且更有可能点击由DIAN推荐的商品。

Conclusion

我们定义了一个新的任务称为微型应用程序内的触发诱导推荐（TIRA），并确定TIRA的关键在于准确估计用户在进入微型应用程序时的意图。为了实现这一目标，我们提出了一种新颖的深度意向感知网络。通过实验证明它在大型工业数据集上的有效性以及在线A/B测试，其效果已在实际应用中得到了验证。

编辑于 2023-11-25 12:04 · IP 属地北京

ctr预估 阿里巴巴集团 工业级推荐系统

赞同 19 2 条评论 分享 喜欢 收藏 申请转载



理性发言，友善互动

2 条评论

默认 最新



土豆小子

方便给个论文链接或者论文名吗？

2023-11-16 · 上海

回复 喜欢



SmartMindAI 作者

arxiv.org/pdf/2211.0865...

2023-11-16 · 北京

回复 喜欢

推荐阅读

NASA计划、阿里巴巴的未来和巨头之间的军备竞赛

作者：范俊杰 喜欢把自己叫做一家数据公司的阿里巴巴，接下来要认真地对待数据、底层科学技术和“人类的未来了”。看起来比Google、Facebook和百度这些对人工智能早有规模布局的公司晚... 硅星人



如何看“低谷”里的阿里？

锦缎研究院



阿里巴巴为未来50年进行的战略布局 终篇

左刀

发表于小商帮

疫情当下，阿里经济体迎机遇

2月13日，阿里巴巴如期在美国公布了2020财年第三财季财报：出意外，阿里巴巴交出了一份收、净利双双高增长且超出预期的成绩单。在去年双十一助推下，阿里巴巴第三财季 太保财评社 发表于创