首页 > 天池大赛 > 赛题与数据

状态 举办方 第1赛季截止日期 总奖池 参赛队

天池新人实战赛之[离线赛]

进行中

詞里云 aliyun.com

2018/05/17

¥ 0

4259

赛制介绍

赛题与数据

FAQ

排行榜 🔐

论坛

提交结果

我的成绩

我的团队

赛题与数据

文件名称	文件格式
选手结果数据样例.csv	.csv (159B) <u></u>
fresh_comp_offline.zip(提取码:e8h6dA)	download 👤

竞赛题目

在真实的业务场景下,我们往往需要对所有商品的一个子集构建个性化推荐模型。在完成这件任务的过程中,我们不仅需要利用用户在这个商上的行为数据,往往还需要利用更丰富的用户行为数据。定义如下的符号:

- U——用户集合
- I——商品全集
- P——商品子集,P ⊆ **I**
- D——用户对商品全集的行为数据集合
- 那么我们的目标是利用D来构造U中用户对P中商品的推荐模型。

数据说明

本场比赛提供20000用户的完整行为数据以及百万级的商品信息。竞赛数据包含两个部分。

第一部分是用户在商品全集上的移动端行为数据(D),表名为tianchi_fresh_comp_train_user_2w,包含如下字段:

字段	字段说明	提取说明
user_id	用户标识	抽样&字段脱敏
item_id	商品标识	字段脱敏
behavior_type	用户对商品的行为类型	包括浏览、收藏、加购物车、购买,对应取值分别是1、2、3、4。
user_geohash	用户位置的空间标识,可以为空	由经纬度通过保密的算法生成
item_category	商品分类标识	字段脱敏
time	行为时间	精确到小时级别

第二个部分是商品子集(P),表名为tianchi_fresh_comp_train_item_2w,包含如下字段:

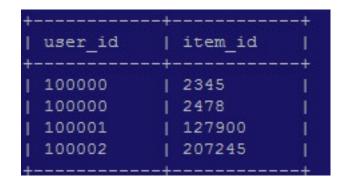
字段	字段说明	提取说明
item_id	商品标识	抽样&字段脱敏
item_ geohash	商品位置的空间标识,可以为空	由经纬度通过保密的算法生成
item_category	商品分类标识	字段脱敏

训练数据包含了抽样出来的一定量用户在一个月时间(11.18~12.18)之内的移动端行为数据(D),评分数据是这些用户在这个一个月之后的

(12.19)对商品子集(P)的购买数据。参赛者要使用训练数据建立推荐模型,并输出用户在接下来一天对商品子集购买行为的预测结果。

评分数据格式

具体计算公式如下:参赛者完成用户对商品子集的购买预测之后,需要将结果放入指定格式的数据表(非分区表)中,要求结果表名为: tianchi mobile recommendation predict.csv,且以utf-8格式编码;包含user id和item id两列(均为string类型),要求去除重复。例如:



评估指标

比赛采用经典的精确度(precision)、召回率(recall)和F1值作为评估指标。具体计算公式如下:

$$\begin{aligned} & \text{Precision} = \frac{\mid \cap (\textit{PredictionSet}, \, \textit{ReferenceSet}) \mid}{\mid \textit{PredictionSet} \mid} \\ & \text{Recall} = \frac{\mid \cap (\textit{PredictionSet}, \, \textit{ReferenceSet}) \mid}{\mid \textit{ReferenceSet} \mid} \\ & \text{F1} = \frac{2 \times \textit{Precision} \times \textit{Recall}}{\textit{Precision} + \textit{Recall}} \end{aligned}$$

其中PredictionSet为算法预测的购买数据集合,ReferenceSet为真实的答案购买数据集合。我们以F1值作为最终的唯一评测标准。

特别感谢:

本次大赛出题方阿里巴巴集团-商家事业部-应用技术部。

该机构致力于通过大数据赋能商家,为所有的商家提供全方位的SaaS应用技术服务。提供官方服务产品包括聚星台(http://juxingtai.taobao.c 阿里官方CRM产品(http://crm.taobao.com)、商家评价管理、商家智能客服。

该机构通过大数据技术帮助商家实现店铺、商品详情、微淘、工作台的千人千面,提升会员运营、流量运营和精准营销的效率; 同时引入服务 并且建立服务市场(http://fuwu.taobao.com),使得商家可以借助服务商实现个性化、精细化的业务运营

关于天池

切换语言

中文 | English

了解更多,请关注天池微信 🤲

阿里巴巴集团 淘宝网 天猫 聚划算 全球速卖通 阿里巴巴国际交易市场 1688 阿里妈妈 阿里旅行•去啊 阿里云计算 YunOS 阿里通信 万网 高德 优视 友盟 酷盘 虾米 天天动听 来 11 Main 支付宝



