



状态

主办方

第 2 赛季截止日期

总奖金

参赛队

商场中精确定位用户所在...

进行中



2017/12/11

¥ 100000

2845

赛制介绍

赛题与数据

FAQ

排行榜

技术圈

文件名称 (报名后可下载)	文件格式
AB检测试集-evaluation_public.csv	.csv (121MB)
训练数据-cdf_first_round_shop_info.csv	.csv (371KB)
训练数据-cdf_first_round_user_shop_behavior.csv	.csv (281MB)

1、数据

本赛题提供两种数据

- 商场内店铺的信息数据，这个对训练和评测都是统一的。
- 真实用户在这些商场内的一段时间的到店交易数据，训练和评测将采用不同的时间段。

注意：为了保护用户和商家的隐私，所有数据均作匿名处理，同时做了必要的有偏采样、过滤等脱敏措施，某些数据的某些字段可能为NULL，请自行处理。

2、评价方式

本赛题采用准确率进行评价，详细评分算法如下：

预测正确：您给出的shop_id和标准答案的shop_id相等。

准确率 = 预测正确样本总数/总样本数

注意：如果某些row_id样本您没有给出结果，那么也会被当成错误识别。

3、字段表

Table 1、店铺和商场信息表（复赛的表名为：ant_tianchi_cdf_sl_shop_info）

Field	Type	Description	Note
shop_id	String	店铺ID	已脱敏
category_id	String	店铺类型ID	共40种左右类型，已脱敏
longitude	Double	店铺位置-经度	已脱敏，但相对距离依然可信
latitude	Double	店铺位置-纬度	已脱敏，但相对距离依然可信
price	Bigint	人均消费指数	从人均消费额脱敏而来，越高表示本店的人均消费额越高
mall_id	String	店铺所在商场ID	已脱敏

Table 2、用户在店铺内交易表（复赛的表名为：ant_tianchi_cdf_sl_user_shop_behavior）

Field	Type	Description	Note
user_id	String	用户ID	已脱敏
shop_id	String	用户所在店铺ID	已脱敏。这里是用户当前所在的店铺，可以做训练的正样本。（此商场的所有其他店铺可以作为训练的负样本）
time_stamp	String	行为时间戳	粒度为10分钟级别。例如：2017-08-06 21:20
longitude	Double	行为发生时位置-经度	已脱敏，但相对距离依然可信
latitude	Double	行为发生时位置-纬度	已脱敏，但相对距离依然可信
wifi_infos	String	行为发生时Wifi环境，包括bssid（wifi唯一识别码），signal（强度），flag（是否连接）	例子：b_63964801-67/false;b_411245141-86/false;b_287233271-90/false;解释：以分号隔开的WiFi列表，对每个WiFi数据包含三项：b_6396480是脱敏后的bssid，-67是signal强度，数值越大表示信号越强，false表示当前用户没有连接此WiFi（true表示连接）。

Table 3、评测集（复赛的表名为：ant_tianchi_cdf_sl_test）

测试数据A榜和B榜格式相同，只是选取的时间不同，A榜数据是9月份第一周数据，B榜数据是9月份第二周数据。

Field	Type	Description	Note
row_id	String	测试数据ID	
user_id	String	用户ID	已脱敏，并和训练数据保持一致
mall_id	String	商场ID	已脱敏，并和训练数据保持一致
time_stamp	String	行为时间戳	粒度为10分钟级别。例如：2017-08-06 21:20
longitude	Double	行为发生时位置-经度	已脱敏，但相对距离依然可信
latitude	Double	行为发生时位置-纬度	已脱敏，但相对距离依然可信
wifi_infos	String	行为发生时Wifi环境，包括bssid（wifi唯一识别码），signal（强度），flag（是否连接）	格式和训练数据中wifi_infos格式相同

Table 4、选手需要提交的结果，统一命名为：result（复赛结果表：ant_tianchi_cdf_sl_predict，需要参赛者自行创建该表并存放最终结果，格式见下文）

注意：每次提交的结果集需要和评测集数据记录数保持一致，避免评测无结果。

Field	Type	Description	Note
row_id	String	测试数据ID	下载的测试文件中的row_id
shop_id	String	店铺ID	算法检测的结果

答案样本

row_id	shop_id
1	xx
4	xx
7	xx
10	xx

注意：在数加平台读取表时，请在表名前加前缀odps_tc_257100_f673506e024，形如：
desc odps_tc_257100_f673506e024.ant_tianchi_cdf_sl_shop_info

关于我们 法律协议 了解更多，请关注天池微信

解决方案

成熟解决方案 | 专业架构师咨询 | 成功客户案例分享 | 专属定制服务

行业架构师咨询

热门产品	云服务器ECS 云计算	云数据库RDS	云存储OSS	NAT网关	负载均衡	域名注册	网站建设	大数据
用户热搜	网站备案 云安全	网安法	CDN加速	API网关	企业邮箱	whois查询	视频直播	视频转码
更多推荐	全民云计算 合规安全解决方案	免费套餐	学生机	IT论坛	数据可视化	云虚拟机	com域名	cn域名

关于我们 法律声明 廉正举报 友情链接

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 飞猪 阿里云计算 YunOS 阿里通信 万网 高德 UC 友盟 阿里星球 来往 钉钉 支付宝

© 2009-2016 Aliyun.com 版权所有 ICP证：浙B2-20080101

浙公网安备 33010002000099号