


现金巴士 | CashBUS

微额借款用户人品预测大赛

DC微信公众号 名称：DataCastle
DC QQ群 群号码：423732457

距离结束还有18天

2015-11-26 00:00:24

算法区

2016-02-29 23:59:59

竞赛告示牌

主页

竞赛数据

提交结果

竞赛信息

竞赛描述

竞赛规则

评价标准

主办方

排行榜

我的队伍

我的提交

讨论区

csv提交tips

提交数据的标签只能是0...

破空神棍惊世现身，各...

你真的了解AUC么？关...

查看全部讨论>>

排行榜

1.不得直视本王

2.云泛天音

3.TNT_000

4.就是gan

5.算得太慢怪我咯

6.sheng123

7.lambda

8.getmax

9.没什么好名字

10.td_ec_am

11.史高治.高柏飞

12.TDAM

13.156

14.326032480

竞赛数据

数据描述

我们为参赛者提供了大量非常宝贵的、来自微额借款行业第一线的实战数据。不仅有常规的带标签数据，还有无标签的数据供大家挑战semi-supervised learning。数据集中共有1138个特征，以用户的多维度行为数据为主。既有数值型特征，也有类别型特征，且均经过脱敏处理。

数据主要包含以下几类：（编码均为UTF-8）

- 1) 训练集（带标签）：15,000个样本
带标签的训练集中共有15,000个样本。train_x.csv中存有样本的特征信息，uid为样本的id，x0、x1、x2...为特征。train_y.csv中存有样本的标签信息，uid为样本的id，y为样本的标签：1为正样本（人品杠杠滴），0为负样本（人品堪忧）；
- 2) 测试集：5,000个样本
test_x.csv中存有测试集的特征信息，格式同train_x.csv。参赛者的目标是尽可能准确地地区分测试集中样本的标签。
- 3) 训练集（无标签）：50,000个样本
在微额借款的真实场景中，除了放款的客户（人品已知），还有相当一部分被拒绝的客户，他们的人品是未知的。为了提高本次比赛的趣味性与挑战性，我们从中挑选了50,000个样本，存在train_unlabeled.csv中，格式同train_x.csv。供参赛者进行semi-supervised learning的探索。
- 4)特征描述：
features_type.csv为本次比赛的1138个特征的类型资料；feature为特征名：x1, x2, x3...type为特征类型：numeric（数值型）或category（类别型）。

数据获取

地址：http://pan.baidu.com/s/1i36Q4p3；特征描述下载：http://pan.baidu.com/s/1bnBgO5h
密码：m5bq；特征描述：hkmi

任务描述

- 1) 构建优秀的分类器；
- 2) 提交分类器对测试集中样本的评分（评分越高，表明样本人品好的可能性越大）。

15.Cauchy_
16.PF_Team
17.johnnylee
18.草上飞
19.Newton
20.100brain
查看全部排行榜>>

支队伍

次提交

关于我们

服务条款
隐私协议
我们的客户
联系我们

商务合作

联系人：周莹
电话：18080138761
邮箱：ying.zhou@hirebigdata.cn

DC QQ群

名称：DataCastle
群号：423732457

DC 微博

名称：DataCastle



DC微信公众号