## 前海征信“好信杯”大数据算法大赛

## 赛题描述：

作为平安旗下专业第三方商业征信机构，前海征信有着丰富的数据资源。

本次赛事中主办方前海征信开放业务数据，设计国内首个迁移学习赛题：参赛选手需依据给定的4万条业务A数据及4千条业务B数据，建立业务B的信用评分模型。其中业务A为信用贷款, 其特征是债务人无需提供[抵押品](http://baike.baidu.com/item/%E6%8A%B5%E6%8A%BC%E5%93%81)，仅凭自己的信誉取得贷款，并以借款人信用程度作为还款保证；业务B为现金贷，即发薪日贷款（payday loan），与一般的消费金融产品相比，现金贷主要具有以下五个特点：额度小、周期短、无抵押、流程快、利率高，这也是与其借贷门槛低的特征相适应的。

由于业务A、B存在关联性，选手如何将业务A的知识迁移到业务B，以此增强业务B的信用评分模型，是本次比赛考察的重点

## 比赛日程

1.报名日期：启动时间：2017-04-05  截止时间：2017-05-10

2.比赛时间：2017-04-05 至 2017-05-10

3.赛事截止日期：2017-05-10

4.初赛评审时间：2017-05-11 至 2017-05-20

5.线下决赛时间：2017-06-03

## 参赛选手要求

1.高等院校、科研单位、互联网企业等人员均可报名参赛

2.本次大赛建议组队参加（团队人数≤5人），也欢迎个人报名参赛

3.本次大赛中，按照决赛成绩，排名前五的学生团队可额外获得奖学金，其中学生团队指的是全体队员皆为学生

4.可以接触到赛题相关数据的人员可以报名参加比赛，但参加比赛作品将不计入最终排行榜参与评奖

## 比赛数据

比赛数据集共包含：A\_train.csv，B\_train.csv，B\_test.csv三个数据文件。

A\_train.csv（feature+label）为40k条，贷款期限为1-3年，平均贷款金额为几千至几万的中等额度信用贷款业务A训练数据。

B\_train.csv(feature+label)为4k条，贷款期限为7-30天，平均贷款金额为一千的小额短期贷款业务B训练数据。

B\_test.csv(feature)为13k条业务B测试数据。

训练集里的feature包括了no, ProductInfo\_1, ProductInfo\_2, ... , UserInfo\_1, UserInfo\_2, ... , WebInfo\_1, WebInfo\_2 ...，以及其他除flag字段外小写字母开头的字段。

其中no是每条记录的id，flag字段即是label。训练集A和训练集B的字段相同。

更多信息参考:**数据说明.docx**

## 评审规则

1.4月5日至5月10日提供每天5次的评测和排名机会，排行榜将进行实时更新，按照评测指标从高到低排序

2.评审使用AUC（Area Under ROC Curve）值来判断分类模型的好坏。[参考链接](http://scikit-learn.org/stable/modules/generated/sklearn.metrics.auc.html)

3.初赛将综合提交结果和迁移学习解决方案的评审表现，选出10支队伍进入决赛。其中提交结果分数占比70%，迁移学习解决方案分数占比30%

4.需线下答辩

## 提交要求

1. 选手提交需要提交由no和pred两列组成的UTF8无bom格式的csv文件，逗号分隔。其中no为测试数据里的no，pred为flag为1概率值。

2. 提交样例文件，请见文件“submit\_sample.csv”

3. 文件命名规则为prediction\_playername\_date，date要求为yyyymmdd形式

4. 提交截止时间：2017-05-10

5. 评审使用AUC（Area Under ROC Curve）值来判断分类模型的好坏。[参考链接](http://scikit-learn.org/stable/modules/generated/sklearn.metrics.auc.html)

6. 得奖选手要求后续需提供实现代码及说明PPT

7. 要求代码为python/R/Java/C格式

## 比赛奖项

为鼓励学生选手创新思维，勇于挑战最新算法，大赛设置**常规奖**和**奖学金**共**8**组奖项。

其中综合排行奖将奖励综合排名第一、二、三名的团队；学生专项奖将奖励综合排名第一至第五的学生团队。综合排名第一的学生团队有望赢取两万元奖金。

本次大赛奖金池共4万元人民币，设置常规奖3项，奖励排名前三的团队；同时设置奖学金5项，额外奖励排名前五的学生团队（指全体队员皆为学生身份）。奖项具体包括：

一等奖   10000元 1支团队

二等奖   6000元  1支团队

三等奖   3000元  1支团队

学生团队（指全体队员皆为学生身份）可额外获得：

一等奖学金   10000元 1支团队

二等奖学金   5000元  1支团队

三等奖学金   2000元  3支团队

## 赛题答疑与反馈

为更好的使大家能通过比赛成长自我，收获良师益友，获得好成绩！我们提供了比赛论坛和QQ群两种选手交流和反馈渠道。由于论坛可以更好的沉淀内容，所以我们建议选手通过论坛进行交流和问题反馈。

1.QQ群：605318706。

2.论坛：比赛详情页下。特别设有答疑汇总贴。