

玩家付费金额预测竞赛代码说明

运行环境

保证电脑内存 16 G 及以上，不过 macOS 8G 的内存也能跑，但是在其他 8G 电脑上会爆内存。以下环境都测试过可以运行，结果会有细微差异。

- macOS 10.13.6
- Kaggle Kernels
- win10 版本 1709

所需依赖包

- anaconda 1.6.5
- Keras 2.1.5
- lightgbm 2.0.11
- numpy 1.14.2
- pandas 0.20.3
- scikit-learn 0.19.0
- scipy 0.19.1
- xgboost 0.6

运行说明

在 `src` 同级目录下建立文件夹 `data`，并且将训练集和测试集解压放在里面；

在 `src` 文件夹里面建立 `output` 文件夹（压缩包里已有可以不建立）；

在终端进入代码路径，然后使用 `bash` 运行 `main.sh` 文件：

```
cd /src
bash -x main.sh
```

最终得到输出文件 `submission.csv`，存放于 `src` 文件夹目录下。

如果 `bash` 文件运行出错，或者不能得到最终输出文件，请按以下步骤，在终端里分别执行：

```
python _1_gbd1_model.py
python _1_gbd2_model.py
python _2_nn_model.py
python _2_gbd_model.py
python _2_knn_lasso_model.py
python _2_lgb_model

python _3_merge.py
```

macOS 和 win10 ,Kaggle Kernels 跑代码 `_2_knn_lasso_model.py` 输出的结果差异较大, 如果输出的预测结果不正常, 请使用 `knn_lasso2.ipynb` 运行一下, 将其输出结果代替

`_2_knn_lasso_model.py` 的结果, 然后重新运行 `_3_merge.py` 得到最终提交结果。

由于比赛时不同队员写的代码风格不同, 有些不规范, 最后提交的是重新整理的代码, 如果复现有任何问题请及时联系我们, 谢谢。

玩家付费金额预测竞赛 Rank 1 团队 `挖矿中`。