

机器学习导论·大作业·“交通拥堵”

机器学习导论·大作业·“交通拥堵”

基本信息

竞赛网址

报名截止时间

大作业截止时间

赛题简介

赛题背景

数据说明

赛题简析

大作业注意事项

组队事宜

队伍命名

队伍名称登记

评分标准

联系助教

基本信息

竞赛网址

<https://competition.huaweicloud.com/information/1000040088/circumstance>

报名截止时间

2020年06月03日18点

大作业截止时间

2020年06月19日

请在上述时间之前，完成竞赛官网上的结果提交，并将报告通过教学立方提交。

赛题简介

赛题背景

本赛题以深圳北站为背景，期望分析出深圳北站周边道路上，车辆对交通拥堵和北站出行的影响，实现基于网约车轨迹数据对路段交通拥堵指数的预测。

数据说明

数据分为两大类：交通拥堵指数数据、网约车轨迹数据，前者中包含字段"time"，后者中包含时间戳。两者设计的时间段是重叠的，两者的转换或许会对模型有帮助，关于时间戳的概念详见：

https://en.wikipedia.org/wiki/Unix_time。

其他数据说明详见竞赛网址。

赛题简析

本题需要对原始数据进行较多的预处理，以得到机器学习模型所需的特征，这一预处理过程会很大程度上决定模型的性能。得到特征之后，即为机器学习中常见的回归问题。

在完成预处理和特征工程后，可以使用网络上已开发好的机器学习工具包，如

- lightgbm <https://github.com/microsoft/LightGBM/tree/master/python-package>
- xgboost https://xgboost.readthedocs.io/en/latest/python/python_intro.html
- pytorch <https://pytorch.org/>

当然，也可以使用不局限于以上推荐的工具包完成赛题。

大作业注意事项

组队事宜

1-3人1组。

队伍命名

请务必使用**符合要求的队伍名**。

队伍名中需要包含下列字母，只需出现即可，不限定次序：

- "n" & "j" & "u"，但尽量不要按"nju"出现
- 所有队员的姓名首字母

如下列各例为张三同学的合法队伍名：

- jintianzaisuzhou
- nojazzmusic

队伍名称登记

队名及队伍成员名单请尽快确定，并QQ提交给助教（周大蔚、吴锦辉），一经提交，不能更改。

评分标准

由打榜成绩与报告成绩综合决定。

联系助教

关于该大作业的疑问，可以咨询助教周大蔚、吴锦辉。