

Rossman 店铺销量预测

2018 年 5 月 27 日

陈俞飞

题目描述：

Rossman 是一个在欧洲 7 个国家有 3000 家分店的药品连锁店。我们需要根据 Rossman 提供的过去 2-3 年的各种店铺信息来预测 Rossman 接下来 6 周的销售额。

数据下载：

此数据集可以从 Kaggle 上下载。

下载地址：<https://www.kaggle.com/c/rossmann-store-sales/data>

开发环境与所需工具：

开发系统：win10 64bit

处理器：I5-8400 CPU 2.8GHz

内存：8G

软件：anaconda, Jupyter Notebook, Python 3.7

库：pandas, numpy, datetime, matplotlib, os, gc, xgboost, operator, seaborn

训练时间：最终的特征筛选完后，进行训练，单词训练时间约为 4 小时左右

项目相关代码文件：

| 序号 | 文件列表 | 文件说明 |
|----|---------------------------|---------|
| 1 | EDA.ipynb | 探索性数据分析 |
| 2 | feature engineering.ipynb | 特征工程 |
| 3 | base model.ipynb | 基准模型 |
| 4 | final model.ipynb | 最终模型 |

最终评分：

| Private Score | Public Score |
|---------------|--------------|
| 0.11709 | 0.11072 |