

Readme

1.1 代码运行环境：

windows10

Python3.6 及以上

依赖库：

Numpy、pandas、LightBGM、matplotlib、warnings、Seaborn、scipy、sklearn、gc、gurobipy

需要安装 licence: <https://zhuanlan.zhihu.com/p/52371462>

1.2 解决方案及算法

1、数据分析以及清洗

删除只有一个值的列，更改预测的 label 为 qty_diff（qty 净增量），箱线图分位点删除异常值；

2、特征工程

日期特征（年、月、日、季度、月份、周、星期天等）；

历史时间点特征（1,2,3,7,14,21 天）；

滑窗统计特征（7、14、21 滑窗）；

日期差特征；

上个月份、季度、周的统计量特征（为了不使用未来数据）；

类别特征标签编码；

单个 geography 和 product 中 unit 的统计量特征；

训练集 unit 关于 qty_diff 的统计量特征。

3、模型

LightBGM（包含缺失值列和类别特征）

4、补货策略

4.1 根据评价指标，用 gurobipy 做优化；

4.2 根据今天补货 14 天后到，7 天补货一次；预测未来 14 天到 21 天的使用量（按照最大子序列索引所在位置，在 MSS 前面的直接加和，后面的如果和小于 0 不加，大于 0 加上）。