

毕业项目开题报告

项目背景

项目选择 Forecast Rossmann Store Sales。这是一个药店销售量的预测项目。数据有 2013-2015 的销售数据。可以根据数据性预测。选择这个项目因为对商业预测比较感兴趣。另外考虑这些预测方法是否可以应用到股票等其他金融工具上。

问题描述

原始数据包括内容有：店 ID，周几，日期，客户数量，是否开业，是否有促销，是否有节假日，是否是学校假日。需要通过这些数据预测销售额。

输入数据

原始数据包括内容有：店 ID，周几，日期，客户数量，是否开业，是否有促销，是否有节假日，是否是学校假日。

解决办法

这个应该本身提供基础数据应该使用监督学习。销售金额这种预测考虑使用回归方式预测，会尝试线性回归，逻辑回归，以及决策树回归，通过测试选择效果的好的回归模型。

基准模型

基准模型考虑采用逻辑回归、线性回归、决策树回归进行测试。

评估指标

通过 kaggle 官网的 RMPSE 进行评分。

设计大纲

初步想法是使用逻辑回归，线性回归，决策树回归与交叉测试进行训练与参数优化，前期去重字段是日期，并对星期数据进行归一化。后面可以根据效果考虑去除不同字段进行优化。并进行最终结果预测与测试。