Rossmann销售预测报告

# 问题的定义

　　这是Udacity机器学习课程的毕业项目，项目中将应用机器学习算法求解销售预测问题。

## 销售预测 – 老问题求新思路

作为销售计划的中心任务之一，销售预测是指通过一定的手段和方法，对未来特定时间内全部产品或者特定产品的销售数量或销售金额的估计。在操作层面上，销售预测是在充分考虑了未来各种影响因素的基础上，结合本企业的过往销售业绩，运用一定的分析方法提出切实可行的销售目标。

影响销售的因素很多，包括有需求变化、经济变动等成分的外界因素，和有营销策略、生产状况、销售人员等成分的内部因素，但销售预测对于完善客户需求管理、指导运营、优化供应链、提高企业利润方面都具有重大促进作用，降低企业的业务计划的不确定性，因此，提出合理的销售预测一直也是人们孜孜不倦的追求，是企业辅助决策的重要工具。销售预测也几乎成为了所有商业智能（BI）的终极目标。

多年来，人们已经形成了定性分析法和定量分析法两类分析方法。其中，定量分析法中的趋势预测分析法和因果预测分析法在实际应用中也能取得一定的预测效果。他们可以基于历史数据，运用一些直观的算法，如算术平均法、指数平滑法，来进行预测，相比定性分析法，预测效果也有了一定的提高[1]。

但随着经济全球化，商业网络化的进程，市场竞争日趋激烈，业务复杂化，数据规模海量化，传统的预测方法已经越来越力不从心。预测的精确性围城销售预测的核心痛点，人们迫切希望一些性能更高，更智能的预测方法。数据挖掘[2]技术由计算机自动从大量数据集中提取隐含的、事先不知道但有潜在应用价值的信息，可用于学习复杂销售的复杂特征对于销售的影响，从而得到较好的预测效果。本课题将应用机器学习算法来实施销售预测。

## 问题分析

Rossmann是欧洲的一家连锁药店，在欧洲7个国家拥有超过3000家药店。这是一个Kaggle[比赛项目](https://www.kaggle.com/c/rossmann-store-sales)，本课题将按照项目中的说明，需要为Rossmann在德国的1115家药店做出提前6个星期的每日销售预测。

对于Rossmann的销售预测问题将是一个具体领域的销售预测问题，作为药店连锁店，具有一定的行业特征，这些将体现为数据特征。如上节所述，基于机器学习算法的数据挖掘技术会是一种可能的得到一个合适预测模型的办法。

本课题得到的预测模型将用于输出未来6个星期里每天的销售量，预测结果可以和实际销售情况对比，从而衡量预测效果。在实际使用中，还可以随着时间的推移，不断学习和预测。

## 评价指标

本项目将使用Kaggle项目中的评估函数RMSPE（Root Mean Square Percentage Error）作为评估指标：

其中，为一个门店某天的销售量，为相应的预测值。

# 研究方法

研究

# 项目结果

# 结论分析

分析一下