家乐福

销售预测与需求计划

项目文档

目录

1. **项目背景与解决方案(**解决方案详细文件**)**
   1. 项目背景
      1. 销售预测
      2. 需求计划
   2. 解决方案
      1. 销售预测模型
      2. 需求计划模型
2. **系统设计与架构(**技术文档**)**
   1. 硬件架构
   2. 工作流程
   3. 数据流程
3. **数据处理与特征工程(**技术文档**)**
   1. 使用的家乐福数据（表名以及列名）
      1. 交易记录数据
      2. 商品信息数据
      3. …
   2. 数据处理目标以及方法
      1. 剔除BP以及非正常销售
      2. 区分DM销售以及非DM销售
      3. …
   3. 数据处理流程图示
   4. 输出表说明
      1. 模型特征信息表
      2. 需求计划信息表
4. **销售预测的建模(**技术文档**)**
   1. 非DM期间的销售预测模型
      1. 模型介绍
      2. 模型训练使用的特征
      3. 模型训练的逻辑
      4. 模型输出的结果表说明
   2. DM期间的销售预测模型
      1. 模型介绍
      2. 模型训练使用的特征
      3. 模型训练的逻辑
      4. 模型输出的结果表说明
5. **需求计划的逻辑(**技术文档**)**
   1. 使用的数据
   2. 运作逻辑
   3. 输出的文件说明
6. **使用说明**(用户文档)
   1. 代码文件夹内容介绍
   2. 执行方法
      1. 代码中的参数设定
      2. 设定自动执行任务（Airflow上）
      3. 在Airflow的操作方法
7. **测试方法**(测试文档)
   1. 样本数据说明
   2. 测试的执行方法
   3. 日志检索以及排错方法
8. **安装和部署方法**(部署文档)
   1. 系统运行环境的要求
   2. 安装和配置
      1. Airflow
      2. Python3 library
      3. Spark
9. **项目背景与解决方案**
   1. 项目背景

数据咨询公司Artefact为家乐福开发一套销售预测与需求计划的模型，旨在提高供应商到货率，降低大仓和门店库存，已提高大仓和门店的商品上架率。此模型分为销售预测模型和需求计划模型两部分。

* + 1. 销售预测模型

Artefact为家乐福开发一个基于机器学习的销售预测模型。此模型将主要食用家乐福数据池历史数据，按周预测华东区49家大型超市3个供应商（宝洁，联合利华和雀巢）的未来9周的销售预测。产品范围涵盖正常商品和海报单品。此项目模型为概念验证型模型。

* + 1. 需求计划模型

Artefact将结合供应商和大仓订单参数，把销售预测转化成大仓和门店的订单。此项目模型为概念验证型模型。

* 1. 解决方案

项目解决方案主要分为两个板块：销售预测模型和需求计划模型。

* + 1. 销售预测模型

销售预测模型主要分为5个步骤：商业数据源获取，数据清理和转化，模型特征转化，模型训练，结果输出。

步骤一 商业数据源获取

在项目前期，Artefact的咨询顾问对家乐福业务团队与IT团队进行访谈，明确影响销售的商业因素，并得知商业数据是如何储存在数据池中。采访记录与采访报告参见交付物3《访谈报告》。

步骤二 数据清理和转化

Artefact的数据工程师对于数据池中的数据库进行处理。具体处理过程参照技术文档。

步骤三 模型特征转化

Artefact的数据科学家将表中数据提炼成DM与非DM模型的特征，具体转化过程参照技术文档。

步骤四 模型训练

模型的特征将会加入到DM和非DM模型中进行模型训练：非DM模型为单周单品模型，即每周每单品一个模型；DM模型为一个模型。具体训练过程参照技术文档。

步骤五 结果输出

模型将会产生销售预测的结果，非DM模型输出结果为单品单店每周的销量预测，DM模型输出结果为单品单店DM档期的销量预测。具体输出过程参照技术文档。

* + 1. 需求计划模型

需求计划模型分为非DM模型和DM模型

* + - 1. 非DM模型

非DM需求计划模型主要分为5个步骤：门店每周非DM销售量分成每天销售量，门店订货参数识别，门店未来库存估计，门店下单量计算，大仓下单量计算，具体处理过程参照技术文档。

步骤一 门店每周非DM销售量分成每天销售量

根据单品单店的历史销售规律将门店每周非DM销售量分解成每天销售量，具体处理过程参照技术文档。

步骤二 门店订货参数识别

把供应链团队提供的订货参数倒入模型中，具体处理过程参照技术文档。

步骤三 门店未来库存估计

根据未来销售预测和现有库存情况，推算每个订货周期以及到货周期的库存情况，具体处理过程参照技术文档。

步骤四 门店下单量计算

根据未来预测的库存情况计算门店下单量，并按订货单位（PCB）转换，具体处理过程参照技术文档。

步骤五 大仓下单量计算

汇总门店下单量到华东区所有门店总下单量。

针对于越库单品，东区总下单量即为所有门店加总下单量。

对于存库单品，大仓库存水平在12-24天区间波动。

具体处理过程参照技术文档。

* + - 1. DM模型

DM需求计划模型主要分为5个步骤：DM期间销售预测总量按天分解，越库单品订货，存库单品A类促销进价（PPP）大仓囤货，大仓发送给门店存库单品75%量，存库单品剩余25%量遵循正常订货逻辑

步骤一DM期间销售总量按天分解

将DM期间销售预测总总量在门店层面按天分解。

步骤二 越库单品订货

针对于越库单品，东区总下单量即为所有门店加总销售总量。

步骤三 存库单品A类促销进价（PPP）大仓囤货

针对于存库单品B类，汇总东区门店所有销售总量即为大仓订货量

针对于存库单品A类，如果促销进价小于等与非促销进价的90%，计算DM档期结束后4周的正常门店销量，加总到此档DM销售量得到订货总量

步骤四 大仓发送给门店存库单品75%量

计算存库单品在DM开档前，大仓第一次推送给门店75%的量。

步骤五 存库单品剩余25%量遵循正常订货逻辑

存库单品剩余25%量遵循正常订货逻辑，DM开档后，有需求的门店订货。如果无需求，则转化成大仓库存。

**系统设计与架构**