

1. 规划

1.1 定义目标 and 需求

1.1.1 目标

构建具有应用意义的机器学习案例。

1.1.2 需求

1. 数据集
2. 算法
3. 模型

1.2 研究和资源收集

1.2.1 研究

方向：

1. Collaborative filtering
 - 适用于了解用户属性（评分，行为，性别，地区），不需要物品属性。
 - 向用户推荐其他相似用户买了什么物品。
2. Content-based filtering
 - 适用于了解物品属性。
 - 向用户推荐除了这个物品，还有其他什么物品可以购买。

限制：

1. 获取数据集时，关于用户信息较难获取，而物品属性较容易取得；
2. 现实中，用户属性（行为，性别，地区）容易取得，物品属性也容易取得。

决定：

1. 使用带有用户属性和物品属性的数据集；
2. 尽量贴合商业意义。

1.2.2 资源

- 学习笔记
- Kaggle数据集
- 公开文档

1.3 制定项目时间表

Done.

2. 设计

2.1 技术框架确定

数据集：

Retailrocket recommender system dataset

算法：

lightFM算法

2.2 设计算法和流程

2.2.1 算法初步设计

已获取数据集能够体现用户的行为和商品分类，考虑利用lightFM算法训练模型

2.2.2 确定数据处理流程

数据获取=>预处理=>EDA=>数据分割=>模型训练

2.2.3 确定机器学习模型

利用lightFM算法训练模型

2.3 设计数据模型

不涉及。

3. 开发

3.1 环境配置

Done.

3.2 编码和实现

See another file.