



**数据挖掘**

**上机实验报告**

|  |  |
| --- | --- |
| 专 业： |  |
| 班 级： |  |
| 学 号： |  |
| 姓 名： |  |

2024年10月21日

## 一、实验目的

1. 掌握Apriori算法的实现过程。

## 二、实验内容

#### 1. 杂货店商品数据集频繁项集挖掘实验

（1）假设最小支持度阈值为0.1，复现Apriori算法并在杂货店商品数据集中**找出所有的频繁项集**，可参考以下流程：

1）找出数据集中的频繁1-项集；

2）根据频繁1-项集生成候选2-项集；

3）判断是否可对候选2-项集进行剪枝操作，然后计算剩余候选2-项集的支持度计数，最后筛选出频繁-2项集；

4）重复以上过程直至候选k-项集为空，最后输出所有的频繁项集。

**提交入口：**

<https://send2me.cn/7RW7WnfB/SnuDSgj7_4U0Mg>

## 三、实验报告评分标准

1. 完成实验（1）（70分）

2. 实验过程与实验结论的记录完整（10分）

3. 表述逻辑清晰（10分）

4. 排版工整自洽，图表标题准确（10分）

## 四、实验过程

## 五、实验结论