**报告**

**学号: 姓名: 班级:**

1. **目标**

练习使用关联规则Apriori算法。完成以下两个任务，并与作业里手动计算的结果进行对比分析。

1. **数据**

课堂收集的选课的真实数据和DBLP数据集

**任务 1**

选择以下某一数据进行关联规则的抽取

|  |  |  |
| --- | --- | --- |
| 序号 | 性别 | 囤货表（食物） |
| 1 | 女 |  |
| 2 | 女 |  |
| 3 | 女 |  |
| 4 | 女 |  |
| 5 | 女 |  |
| 6 | 男 |  |
| 7 | 男 |  |
| 8 | 男 |  |
| 9 | 男 |  |
| 10 | 男 |  |

执行Apriori算法，记录算法设置和结果，要求：

1. 给出算法过程，记录参数设置。
2. 分析结果，找出不同频繁项数量（2-4）的关联规则结果。并给出相应的算法设置。
3. 要求给出重要的算法，过程截图，和必要的文字分析。

**任务2**

使用DBLP数据集，提出一种方法，挖掘密切相关的（即经常一起合写文章）合著者关系。