**实验八 关联分析算法实现-Apriori**

【实验目的】

1 理解Apriori关联分析算法的基本原理。

2 掌握Apriori算法的基本编程实现方法。

3 能够对算法相关环节进行一定的优化。

【实验类型】

设计型

【实验学时】

2学时

【实验环境】

Windows 7以上操作系统

Python3.0以上版本

Pycharm开发环境

Spyder开发环境

【实验要求】

Apriori是最常见的关联分析算法之一，其基本步骤是：

（1）令 k=1，生成所有长度为1的频繁集

重复下列步骤，直到不能确定新的频繁集

（2）根据长度为k的频繁集生成长度为k+1的频繁集

（3）修剪掉存在k长度的子集不是频繁集的候选集

（4）扫描所有事务计算每个候选集的支持度

（5）排除不频繁的候选集，仅保留频繁的

**要求：**

**（1）认真参考教程中对Apriori算法的描述了解其候选集的生成方法和规则的生成方法；**

**（2）使用python编程实现算法（生成频繁集候选时，可简单使用union方法，生成规则时，可考虑使用itertools包提供的combinations辅助）；**

**（3）尽力对频繁集的生成和规则的生成算法进行一些优化。**

注意：编码中可考虑使用itertools包，即需要引入：

