

《数据挖掘》

实验报告本

|  |  |
| --- | --- |
| 班 级： |  |
| 学 号： |  |
| 姓 名： |  |
| 指导教师： | **张静** |

信息科学与工程学院

2019年 4月

**实验一 Apriori算法研究与实现**

**一、实验目的**

通过编程实现Apriori算法，并在给定数据集上进行测试，完成频繁模式的挖掘。

**二、实验设备和环境**

**（自己填写）**

**三、实验内容**

编程实现Apriori算法，并在数据集T10I4D1K.dat、T10I4D10K.dat、T10I4D50K.dat、T10I4D100K.dat上完成测试，找到支持度计数分别为：600，800, 1000时，该数据集中所有的频繁项集，并保存在文档中。

**四、实验过程、结果及分析**

1. 实验过程，关键代码，实验结果及分析。
2. 实验中遇到的问题及解决办法。
3. 实验还存在哪些问题。
4. 本实验有哪些收获和心得体会。、
5. 其它需要补充的问题。

（以上内容不少于1000字）

**实验二 基于SVM的数据分类**

**一、实验目的**

通过编程调试实现SVM算法并在给定数据集上进行测试，实现数据分类。

**二、实验设备和环境**

**（自己填写）**

**三、实验内容**

自己编写或从网上下载SVM的源代码，调试运行SVM算法，研究其参数设定，并在给定数据集上完成测试，实现数据分类，并计算准确率（Precision）和查全率（Recall）。

**四、实验过程、结果及分析**

1. 实验过程，关键代码，实验结果及分析。
2. 实验中遇到的问题及解决办法。
3. 实验还存在哪些问题。
4. 本实验有哪些收获和心得体会。
5. 其它需要补充的问题。

（以上内容不少于1000字）

**实验三 基于划分的聚类算法研究与实现**

**一、实验目的**

通过编程实现k均值和k中心点聚类算法。

**二、实验设备和环境**

**（自己填写）**

**三、实验内容**

编程实现k均值及k中心点聚类算法，并在数据集上完成测试，分别给出k=5、10、20时的聚类结果，在实验结果的基础上分析k均值及k中心点算法的优缺点。

**四、实验过程、结果及分析**

1. 实验过程，关键代码，实验结果及分析。
2. 实验中遇到的问题及解决办法。
3. 实验还存在哪些问题。
4. 本实验有哪些收获和心得体会。
5. 其它需要补充的问题。

（以上内容不少于1000字）

**实验成绩：\_\_\_\_\_\_\_\_\_\_\_\_ 教师签名：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**