##### QG工作室数据挖掘小组实验报告

|  |  |  |
| --- | --- | --- |
| 实习生：李炜乐 | 导师：张润鹏 | 日期：2019 年 7月 21日 |

|  |
| --- |
| **实验名称：Apriori关联算法** |
| **已完成内容：**   1. **Apriori算法代码的编写** 2. **将算法运用在mushroom数据集上，并且进行分析** |
| **未完成内容：暂无** |
| **未完成原因：暂无** |
| **需要帮助：暂无** |

|  |  |
| --- | --- |
| **实验总结** | |
| **知识点总结：**   1. **Apriori包括两个大步骤。其中为（1）找到所有的频繁项集（2）根据频繁项集产生强关联规则。** 2. **在第一步找出所有的频繁项集的工作原理是定义一个最小支持度，然后从每个项遍历取**一个，二个，三个…对他们进行排列组合，如果经过排列组合的项的支持度满足最小支持度，则保留下来，否则则舍去。知道没有排列组合满足最小支持度为止。**** 3. ****根据频繁项集产生强关联规则是定义一个最小置信度，然后筛选其中的不足的关联规则。**** | |
| **遇到问题**：数据集中存在很多相重的字符，却属于各种不一样的特征，这样会导致结果的改变 | **解决过程：将各种数据集的各种特征采取用数字的代替的方式** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **导师评价** | | | | |
| **实验分数** | **知识掌握情况** | **代码编写能力** | **建议** | **评价日期** |
|  |  |  |  |  |