|  |
| --- |
|  |
| —《数据挖掘应用》考核方案 |

2019年11月

一、考核制度

1、该门课程以课程标准规定的课程目标、教学内容和要求为评价依据,采用大作业考核方式实施考核。

2、学生的最终成绩由学生的平时成绩(40%)、大作业(60%)的成绩之和。

3、在平时的课程训练中，学生必须完成每个项目设定的任务，平时成绩为：出勤、学习态度、课堂纪律、平时作业、课堂提问等五项成绩。

4、大作业成绩为现场大作业的得分。

二、大作业考核

**基于XXXX算法的数据分析与挖掘，**完成如下作业：

1．背景和挖掘目标；（10分）

2．数据探索分析代码和结果分析；(包含数据搜集)（20分）

3．数据预处理代码和结果分析；（20分）

4. 数据模型构建代码和结果分析；（30分）

5. 模型进行评价的代码和结果分析；（20分）

**以“基于决策树算法的电力窃漏电用户的数据分析”为例**

（1）背景与挖掘目标

（2）实现数据探索分析的代码和结果分析；

（3）实现数据预处理的代码和结果分析

（4）实现基于决策树的电力窃漏电用户的模型构建的代码和结果分析

（5）对构建的模型进行评价的代码和结果分析

三、参考选题（以下选题可以参考，也可以自行选题）

|  |  |  |
| --- | --- | --- |
| 序号 | 算法 | 选题 |
| 1 | 关联规则算法 | 超市客户购买产品的分析 |
| 2 | 餐饮客户点菜菜品的分析 |
| 3 | 中医证型关联规则挖掘 |
| 4 | 淘宝商品相关推荐的分析 |
| 5 | 决策树算法 | 别克汽车公司的客户购买行为的分析 |
| 6 | 上网用户婚姻情况分析与识别 |
| 7 | 保险公司客户购买行为的分析 |
| 8 | 家电用户行为分析与事件识别 |
| 9 | 聚类算法 | 电信公司客户的细分 |
| 10 | 汽车4S店客户价值的分析 |
| 11 | 网游客户价值的分析 |
| 12 | 城镇居民消费结构的分析 |
| 13 | 基于基站定位数据的商圈分析 |
| 14 | 线性回归算法 | 滁州房产销售预测分析 |
| 15 | 汽车4S店销售预测分析 |
| 16 | 水色图像的水质评价 |
| 17 | 双11淘宝销售预测分析 |

四、提交作业格式：

“学号姓名数据挖掘应用大作业”文件夹，包含3个文件：

1. 代码文件：**“学号姓名大作业.py”或 “学号姓名大作业.ipynb”**

2. 大作业报告： **“学号姓名大作业报告.docx”**

**3. 数据文件：“XXX.csv”或“XXX.xlsx”**