**《资料仓储、勘探与商业应用》课程教学大纲**

|  |  |
| --- | --- |
| **课程名称：**资料仓储、勘探与商业应用 | **课程类别（必修/选修）：必修** |
| **课程英文名称：**  |
| **总学时/周学时/学分：**54/3/3 | **其中实验（实训、讨论等）学时：21** |
| **先修课程： 程序设计基础** |  |
| **授课时间：1至18周，星期三1-3节** | **授课地点：综合实验楼605机房** |
| **授课对象： 2017跨境电商** |
| **开课院系：**  |
| **任课教师姓名/职称：** 叶贤良/讲师 |
| **联系电话：13712632399** | **Email:120372782@qq.com** |
| **答疑时间、地点与方式：** |
| **课程考核方式：**开卷**（ ）** 闭卷**（ √）** 课程论文**（ ）** 其它**（ ）** |
| **使用教材：《数据挖掘概念与技术》，范明译，机械工业出版社，2007年。****教学参考资料： 《数据挖掘：方法与应用》，徐华编，清华大学出版社，2014年** |
| **课程简介：**《数据挖掘：方法与应用》是计算机相关专业教学计划中一门专业课，是一门汇集统计学、机器学习、数据库、人工智能等学科内容的新兴的交叉性学科，是在信息技术领域迅速兴起的计算机技术。在很多重要的应用领域，数据挖掘都发挥着积极的作用。因此这门课程是计算机相关专业的重要课程之一。本课程深入探讨数据挖掘原理，把信息科学、计算科学和统计学对数据挖掘的贡献融合在一起，培养计算机相关专业高年级本科学生具备初步的科研能力和创造能力 |
| **课程教学目标****1.** 了解数据挖掘技术最新进展和前沿发成果，激发学生兴趣。**2.** 理解和掌握数据预处理方法、包括数据清理、数据规约和数据变换等。**3.** 理解和掌握数据仓库的概念和构建方法，多维数据模型的建立等。**4**. 掌握挖掘大型数据库的关联规则方法，贝叶斯分类技术、K均值聚类方法等。 | **本课程与学生核心能力培养之间的关联（可多选）：****□熟练数据挖掘的思路和技巧****□熟练掌握数据预处理的方法和技巧****□熟悉和掌握大型数据库的挖掘方法和技巧**  |
| **理论教学进程表** |
| **周次** | **教学主题** | **教学时长** | **教学的重点与难点** | **教学方式** | **作业安排** |
| 1 | 绪论 | 3 | 重点：数据挖掘的概念及应用领域难点：数据挖掘的基本步骤、主要技术和应用价值，如何激发学生对数据挖掘应用的领域及成功案例的兴趣 | 课堂讲授 |  |
| 2、3 | 数据预处理 | 6 | 重点：数据的基本处理方法难点：数据处理方法的应用。确保学生对处理方法的理论尚能掌握，进一步增加实际应用方面的操作 | 课堂讲授 |  |
| 5 | 数据仓库 | 3 | 重点：数据仓库的定义、多维数据模型和OLAP技术难点：数据仓库的存储结构模式，概念分层的理解，以及数据仓库的设计，要确保学生能较好地掌握，还需课后多查看一些相关的资料。 | 课堂讲授 |  |
| 6 | 相关性与关联规则 | 3 | 重点：关联规则的挖掘难点：Apriori算法的应用 | 课堂讲授 |  |
| 7 | 相关性与关联规则 | 3 | 重点：关联规则的挖掘难点：Apriori算法的应用 | 课堂讲授 |  |
| 10 | 分类和预测 | 3 | 重点：分类和预测方法的概念难点：几种主要的分类算法的应用和实现 | 课堂讲授 |  |
| 11 | 分类和预测 | 3 | 重点：分类和预测方法的概念难点：几种主要的分类算法的应用和实现 | 课堂讲授 |  |
| 14 | 聚类分析 | 3 | 重点：聚类分析的概念和主要算法难点：几种主要的聚类分析算法的应用和实现 | 课堂讲授 |  |
| 15 | 聚类分析 | 3 | 重点：聚类分析的概念和主要算法难点：几种主要的聚类分析算法的应用和实现 | 课堂讲授 |  |
| 18 | 复习 | 3 |  | 课堂讲授 |  |
| **合计：** | 33 |  |  |  |
| **实践教学进程表** |
| **周次** | **实验项目名称** | **学时** | **重点与难点** | **项目类型（验证/综合/设计）** | **教学****方式** |  |
| 4 | 熟悉R软件的安装和编程思想 | 3 | 利用R中的数学函数处理简单数据的操作 | **综合** | 上机操作 |  |
| 8 | 关联规则挖掘算法 | 3 | 熟悉Apriori算法的应用 | **综合** | 上机操作 |  |
| 9 | 关联规则挖掘算法 | 3 | 决策树归纳、提取分类规则、熟悉C4.5算法的应用 | **综合** | 上机操作 |  |
| 12 | 决策树分类C4.5算法 | 3 | 熟悉NB算法的应用 | **综合** | 上机操作 |  |
| 13 | 决策树分类C4.5算法 | 3 | SVM算法的基本大批量及应用 | **综合** | 上机操作 |  |
| 16 | K-means算法 | 3 | K-means算法的应用 | **综合** | 上机操作 |  |
| 17 | K-means算法 | 3 | K-means算法的应用 | **综合** | 上机操作 |  |
| 合计： | 21 |  |  |  |  |
| **成绩评定方法及标准** |
| **考核内容** | **评价标准** | **权重** |
| 平时考勤 | 根据平时厜勤的情况评定分数 | 0.1 |
| 平时实验 | 根据实验报告情况评定分数 | 0.3 |
| 期末考核 | 根据评分标准评定分数 | 0.6 |
| **大纲编写时间：2019-3-17** |
| **系（专业）课程委员会审查意见：**我系（专业）课程委员会已对本课程教学大纲进行了审查，同意执行。系（专业）课程委员会主任签名： 日期： 年 月 日 |

**注：1、课程教学目标：请精炼概括3-5条目标，并注明每条目标所要求的学习目标层次（理解、运用、分析、综合和评价）。本课程教学目标须与授课对象的专业培养目标有一定的对应关系**

 **2、学生核心能力即毕业要求或培养要求，请任课教师从授课对象人才培养方案中对应部分复制（http://jwc.dgut.edu.cn/）**

 **3、教学方式可选：课堂讲授/小组讨论/实验/实训**

 **4、若课程无理论教学环节或无实践教学环节，可将相应的教学进度表删掉。**