《大数据分析》 课程教学大纲

**表 1 课程基本信息**

|  |  |  |  |
| --- | --- | --- | --- |
| **课程代码** | 18116011 | **课程名称** | 大数据分析 |
| **英文名称** | Bigdata Analysis | | |
| **学分/学时：** 3 学分/48 学时 （理论学时： 28， 实践学时： 20） | | | |
| **适用专业** | 信息管理与信息系统 | **课程类型** | 学科与专业基础必修 |
| **开课单位** | 中原大学 | **开课学期** | 五 |

**一 、课程简介**

大数据分析挖掘是高级数据处理和分析技术 。本课程包括数据挖掘的理论和实践两部分 。本课 程是专业课， 属于高年级专业技术基础应用课程， 注重理论和编程实践及应用结合， 强调理论和应 用性， 同时要把数据分析技术应用到不同场景和领域， 需要跨领域知识， 是一门综合性较强， 应用 类课程。

**二 、课程目标**

**课程目标** **1：** 使学生了解大数据分析和挖掘方法的思想和技术

**课程目标** **2：** 掌握重要数据挖掘技术的基本原理， 并能熟练使用 Python 语言实现主流的数据挖 掘算法

**三 、课程目标与所支撑毕业要求的对应关系**

**表 2 课程目标与所支撑毕业要求的对应关系**

|  |  |  |
| --- | --- | --- |
| **课程目标** | **具体要求** | **支撑的毕业要求** |
| 课程目标 1 | 了解大数据分析的基本 知识 | 毕业要求 3.3 能够在设计环节中体现 创新意识， 考虑社会、健康、安全、 法律、文化以及环境等因素。 |
| 课程目标 2 | 掌握数据挖掘技术及方 法 | 毕业要求 4.1 能比较和选择研究路 线， 独立设计实验方案。  4.2 能正确采集、整理实验数据， 对 实验结果进行关联及建模。 |

**四 、教学内容**

**表 3 教学内容与学时分配**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **章节/单** **元** | **推荐授 课形式** | **教学内容** | **建议** **学时** | **课程思政元素** | **对应的课程目标** |
| 第一章 Python 基本语 法元素 | 课堂讲授为主， 结合网络课程辅 助， 课堂讨论 ， 编程实践等手段 提高教学效果。 | 1.程序设计基本方法 。掌握  IPO，输入-处理-输出的程序设 计方法。  2.Python 开发环境配置。会搭 建 windows 和 linux 等多系统 下 Python 开发环境。  3.简单 Python 程序实例：温度 转换  4.Python 程序语法元素分析 | 3 | 能力培养 | 课程目标 1 |
| 第二章 基本数 据类型 | 课堂讲授为主， 结合网络课程辅 助， 课堂讨论 ， 编程实践等手段 提高教学效果。 | 数字类型及操作， 程序实例 。 字符串类型及操作，程序实例。  Time 库使用。 | 3 | 能力培养、实 事求是、科学 精神、团队协 作精神 | 课程目标 1 |
| 第三章  函数和  代码复  用 | 课堂讲授为主， 结合网络课程辅 助， 课堂讨论 ， 编程实践等手段 提高教学效果。 | 函数定义与使用， 实例七段数 码管绘制 。代码复用与函数递 归， Pyinstaller 打包库使用， 科赫雪花小包裹实例。 | 3 | 能力培养、实 事求是、科学 精神、团队协 作精神 | 课程目标 1 |
| 第四章 组合数 据类型 | 课堂讲授为主， 结合网络课程辅 助， 课堂讨论 ， 编程实践等手段 提高教学效果。 | 集合数据类型及操作， 序列类 型及操作， 基本统计值计算程 序实例， 字典类型及操作 ， jieba 库的使用， 文本词频统 计实例。 | 3 | 能力培养、实 事求是、科学 精神、团队协 作精神 | 课程目标 1 |
| 第五章 数据挖 掘引言 | 课堂讲授 | 数据挖掘产生的背景， 定义 ， 和基本流程 。数据挖掘的主要 方法 。数据挖掘的应用场景与 案例。 | 2 | 价值塑造 | 课程目标 1 |
| 第六章  认识数  据 | 课堂讲授为主， 结合网络课程辅 助， 课堂讨论 ， 编程实践等手段 提高教学效果。 | 数据对象与属性类型， 数据基 本统计描述方法， 数据可视化 方法， 数据相似性和相异性度 量方法。 | 3 | 能力培养、实 事求是、科学 精神、团队协 作精神 | 课程目标 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **章节/单** **元** | **推荐授 课形式** | **教学内容** | **建议** **学时** | **课程思政元素** | **对应的课程目标** |
| 第七章  数据预  处理 | 课堂讲授为主，  辅以课堂讨论及 编程实践 | 数据预处理的基本思想和数据 规范化， 数据离散化， 数据清 洗， 特征提取与特征选择。 | 5 | 能力培养、实 事求是、科学 精神、团队协 作精神 | 课程目标 2 |
| 第八章 关联规 则挖掘 | 课堂讲授为主，  辅以课堂讨论及 编程实践 | 关联规则挖掘背景， 概念， 频 繁模式与关联规则概念， 频繁 项集的典型挖掘方法， 关联规 则的生成方法， 关联规则的其 他类型，关联规则模型的度量。 | 6 | 能力培养、实 事求是、科学 精神、团队协 作精神 | 课程目标 2 |
| 第九章  数据分  类 | 课堂讲授为主，  辅以课堂讨论及 编程实践 | 分类的概念， 决策树分类， 朴 素贝叶斯分类， K 近邻分类 ， 分类性能的度量方法。 | 6 | 能力培养、实 事求是、科学 精神、团队协 作精神 | 课程目标 2 |
| 第十章  数值预  测 | 课堂讲授为主，  辅以课堂讨论及 编程实践 | 数值预测概念， 回归方法， 回 归树与模型树， K 近邻数值预 测和预测误差的度量。 | 4 | 能力培养、实 事求是、科学 精神、团队协 作精神 | 课程目标 2 |
| 第十一 章 聚 类分析 | 课堂讲授为主，  辅以课堂讨论及 编程实践 | 介绍聚类分析概念， 相似度度 量方法， K 均值聚类法， 层次 聚类方法， DBSCAN 方法， 聚类 效果衡量方法。 | 6 | 能力培养、实 事求是、科学 精神、团队协 作精神 | 课程目标 2 |
| 第十二 章 挖 掘复杂 的数据 类型 | 课堂讲授为主，  辅以课堂讨论及 编程实践 | 复杂数据类型挖掘方法， 包括 序列类型， 网络类型， 时间空 间数据， 地理数据， 移动对象 和物联网系统数据， 多媒体数 据， 文本数据， Web 数据， 流 数据等。 | 4 | 能力培养、实 事求是、科学 精神、团队协 作精神 | 课程目标 2 |

**五 、教学方式方法**

课堂讲授为主， 结合网络课程辅助， 课堂讨论， 调查研究， 案例分析等手段提高教学效果。

**六 、考核要求与评价**

**1.考核方式和要求。**

**表 4 考核方式和要求**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **考核方式** | **考核内容** | **百分比** | **考核标准** | **对应的课程 目标** |
| 平时成绩 | 实验 | 40 | 课堂表现及实践 | 课程目标 1 |
| 期末考核 | 期末报告 | 60 | 完成指定项目，依据提交时效及内容 完成度考核 | 课程目标 2 |
| 小计 | | 100 |  | |

平时成绩由： 考勤， 课堂表现， 实验， 分享， 小组贡献构成

期末考核采用课程报告形式， 主要是指定大数据分析的行业应用项目， 完成指定项目设计， 分 析， 与结果展示， 提交文档和可执行带注释代码， 及分析结果 。文档描述清晰， 代码运行正确， 模 型效果性能较好都是相应的踩分点。

**2.课程评价**

课程目标达成评价以授课完成时为时间点， 由任课教师完成 。每学年课程结束后由任课教师根 据教学各环节的学生完成情况进行课程目标达成评价， 评价结果用于持续改进。

**七 、先修课程**

统计学， 算法与数据结构， 程序设计

**八 、推荐教材与参考资料 推荐教材：**

《Python 数据分析与挖掘实践（第二版）》， 张良均， 机械工业出版社， 2019.3

**九 、相关实践教学内容说明**

**表 3 课程实验项目与内容设计**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **序号** | **项** **目** | **内容和要求** | **实验学** **时** | **主要 仪器 设备** | **备注** |
| 1． 1 | 2． Python 基 本语法元 素 | 3． 掌握基本 Python 语法元素， 包括 程序基本结构， 注释， 数据类型， 语句， 函数等概念。  4． 通过学习简单程序例子， 可以编写 简单 python 程序。 | 1 | **PC** |  |
| 5． 2 | 6． 基本数据 类型 | 掌握基本 python 数据类型， 并能在程 序中灵活应用。 | 1 | **PC** |  |
| 7． 3 | 8． 函数和代 码复用 | 掌握 Python 函数与代码复用， 可编程 应用。 | 1 | **PC** |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **序号** | **项** **目** | **内容和要求** | **实验学** **时** | **主要 仪器 设备** | **备注** |
| 9． 4 | 10． 组合数据 类型 | 掌握 Python 组合数据类型， 可编程应 用。 | 2 | **PC** |  |
| 11． 5 | 12． 认识数据 | 介绍数据探索的背景， 概念， 掌握数 据探索的主要方法， 步骤。 | 1 | **PC** |  |
| 13． 6 | 14． 数据预处 理 | 掌握数据预处理的主要方法。 | 2 | **PC** |  |
| 15． 7 | 16． 关联规则 挖掘 | 掌握数据挖掘经典方法， 关联规则挖 掘方法及应用。 | 2 | **PC** |  |
| 17． 8 | 18． 数据分类 | 掌握数据挖掘经典方法， 数据分类方 法及应用。 | 2 | **PC** |  |
| 19． 9 | 20． 数值预测 | 掌握数据挖掘经典方法， 数据预测方 法及应用。 | 2 | **PC** |  |
| 21． 1 0 | 22． 聚类分析 | 掌握数据挖掘经典方法， 聚类分析方 法及应用。 | 2 | **PC** |  |

**大纲编制人： 杨玉文**

**大纲审核人： 郑积仕**

**制订日期： 2024 年 1 月**