**《大数据存储》教学大纲**

**一、课程基本信息**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **课程类别** | 专业课程 | | **课程性质** | 必修 | **课程属性** | 理论 | |
| **课程名称** | 大数据存储 | | | **课程英文名称** | Big Data Storage | | |
| **课程编码** | H35B117F | | | **适用专业** | 数据科学与大数据 | | |
| **考核方式** | 考查 | | | **先修课程** | 数据库系统原理 | | |
| **总学时** | 48 | **学分** | | 3 | **理论学时** | | 28 |
| **实验学时/实训学时/ 实践学时/上机学时** | | | | 上机学时：20 | | | |
| **开课单位** | | | | 人工智能学院 | | | |

**二、课程简介**

《大数据存储》课程意在普及大数据知识，帮助学生熟悉大数据时代的现实意义，掌握大数据存储与管理技术，积极投身于大数据应用。

在大数据生态系统中，基础设施主要负责数据存储以及处理公司掌握的海量数据，应用程序则是人类和计算机系统从数据中获知关键信息的工具。为了应对大数据时代，需要从根本上考虑用于数据存储和处理的平台。

关系数据库和NoSQL数据库是数据库演化过程中的两个里程碑。不同的应用程序需要使用不同类型的数据库，而这恰恰是数据管理系统在过去儿十年间不断发展的动力所在。

**三、课程教学目标**

|  |  |  |  |
| --- | --- | --- | --- |
| **课程教学目标** | | **支撑毕业要求指标点** | **支撑毕业要求** |
| **知**  **识**  **目**  **标** | **目标1：**  系统掌握关系型、非关系型数据库的基本原理；掌握键值数据库、文档数据库、列族数据库的操作和设计；掌握NoSql数据的软件应用，能根据实际需求选择适配的数据库 | 1.1 具备大数据专业基础理论知识和专业知识  1.2 理解大数据专业及相关学科的基本概念、知识结构、典型方法等，建立数字化、算法、模块化与层次化等核心专业意识。 | 1．工程知识 |
| **能**  **力**  **目**  **标** | **目标2：**  在实践中理解非关系型数据库基本知识，培养非关系型数据库的操作能力，并根据实际问题，逐步分析与选用、应用适配数据库的思想和方法；  培养学生发现问题、分析问题和解决问题的能力；  培养学生把现实事物抽象地表示为计算机中的数据的能力；  培养学生认真观察和进行需求分析的能力；  培养学生理论联系实际，将所学知识应用于实践的能力；  培养学生掌握文献检索、资料查询及运用现代信息技术获取相关信息的能力。 | 2.1 具备大数据知识运用能力，能将大数据和云计算基本知识，用于分析和解决复杂工程问题；  4.1 能够基于科学原理并采用科学方法对大数据相关领域复杂工程问题进行研究，包括设计实验、分析和解释数据等，并通过信息综合得到合理有效的结论。  5.2开发、选择与使用恰当的技术、资源、现代工程工具和信息技术工具等，包括对复杂工程问题的预测和模拟，并能够理解其局限性。 | 2．问题分析  4．研究  5．使用现代工具 |
| **素**  **质**  **目**  **标** | **目标3：**  通过本课程的学习，培养作为一个工程技术人员必须具备的坚持不懈的学习精神，严谨治学的科学态度和积极向上的价值观，为未来的学习、工作和生活奠定良好的基础；  培养学生遵纪守法，树立科学的世界观、人生观、价值观和社会主义荣辱观；  培养学生了解专业相关的职业和行业的政策和法律、法规，具有较强的法制意识和素养；  培养学生自主学习和终身学习的理念；  培养学生具备良好的科学态度、团队协作精神和创新意识；  培养学生对个人职业发展方向的认识和思考，以及适应社会和行业发展能力。 | 6.1能够基于工程相关背景知识进行合理分析、评价大数据工程实践和复杂工程问题解决方案对社会、健康、安全、法律以及文化的影响，并理解应承担的责任。  8.2能够在数据科学与大数据技术实践中理解并遵守工程职业道德和规范，履行责任。  9.1具有良好的组织管理、人际交往和团队协作能力。  10.1 具有宽阔的国际视野和跨文化交流、竞争和合作能力。  12.1 具备不断学习和适应发展的能力，能够运用现代信息技术获取相关信息和新技术、新知识，持续提高自己。 | 1. 工程与社会   8. 职业规范  9. 个人和团队  10.沟通  12.终身学习 |

**四、课程主要教学内容、学时安排及教学策略**

**（一）理论教学**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **教学模块** | **学时** | **主要教学内容与策略** | **学习任务安排** | **支撑课程目标** |
| 大数据存储基础 | 4 | **重点：**大数据概念；分布式系统；Hadoop分布式处理基础；大数据数据处理基础；大数据存储技术路线  **难点：**分布式系统；hadoop分布式处理基础  **思政元素：**通过对分布式技术的产生背景及面临的问题引入，让学生树立正确的科学观、人生观、价值观和辩证唯物主义认识。  **教学方法与策略：**线下教学。对于概念、原理课堂上予以讲授。课堂运用讲授法和案例法开展教学，辅以启发式互动拓宽学生学习思路 | 课前：预习大数据分布式知识点  课堂：认真听讲，参与课堂讨论  课后：完成相应练习 | 目标1  目标2 |
| 数据管理技术发展 | 4 | **重点：**数据管理系统发展；关系型数据库；NoSQL发展  **难点：** RDBMS结构；关系型数据库及其局限  **思政元素：**通过引入关系型数据库局限性的案例教学，培养学生的科学探索的精神  **教学方法与策略：**线下教学。对于概念、原理课堂上予以讲授。课堂运用讲授法和案例法开展教学，辅以启发式互动拓宽学生学习思路 | 课前：预习关系型数据库的基础知识  课堂：参与课堂讨论  课后：完成相关练习 | 目标1  目标2  目标3 |
| NoSQL数据模型 | 4 | **重点：**NoSQL数据库的性质；NoSQL数据库实例介绍  **难点：** NoSQL数据库性质  **思政元素：**通过介绍相关案例，让学生了解到选择适配数据库的重要性，培养学生的职业素质和道德规范  **教学方法与策略：**线下教学。对于概念、原理课堂上予以讲授。课堂运用讲授法和案例法开展教学，辅以启发式互动拓宽学生学习思路 | 课前：预习NoSQL的基础知识  课堂：参与课堂讨论  课后：完成相关练习 | 目标1  目标2 |
| 键值数据库 | 4 | **重点：**键值数据库的特征；键值数据库的设计与局限  **难点：** 键值数据库的设计  **教学方法与策略：**线下教学。对于概念、原理课堂上予以讲授。课堂运用讲授法和案例法开展教学，辅以启发式互动拓宽学生学习思路 | 课前：键值数据库预习  课堂：参与课堂讨论  课后：完成相关练习 | 目标1  目标2 |
| 文档数据库 | 6 | **重点：**文档数据库数据结构、基本操作；文档数据库设计；MongoDB  **难点：**文档数据库基本操作、设计  **教学方法与策略：**线下教学。对于概念、原理课堂上予以讲授。课堂运用讲授法和案例法开展教学，辅以启发式互动拓宽学生学习思路 | 课前：文档数据库知识预习  课堂：参与课堂讨论  课后：完成相关练习 | 目标1  目标2 |
| 列族数据库 | 6 | **重点：**列族数据库架构、组件；列族数据库设计；HBase  **难点：**列族数据库设计  **教学方法与策略：**线下教学。对于概念、原理课堂上予以讲授。课堂运用讲授法和案例法开展教学，辅以启发式互动拓宽学生学习思路 | 课前：列族数据库知识预习  课堂：参与课堂讨论  课后：完成相关练习 | 目标1  目标2 |

**（二）实践教学**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **实践类型** | **项目名称** | **学时** | **主要教学内容** | **项目**  **类型** | **项目**  **要求** | **支撑课程目标** |
| 上机 | RDBMS与SQL | 4 | 重点：关系型数据库操作、设计  难点：关系型数据库设计  思政元素：开拓创新、勇于探索、积极竞赛，对立积极向上、敬业、诚信、友善的就业风貌，践行社会主义核心价值观的行为准则 | 综合 | 1.掌握关系型数据库基本操作  2.了解的基本关系型数据的设计 | 目标1  目标3 |
| 上机 | 键值数据库 | 4 | 重点：Redis键值数据库安装；键值数据库设计、操作  难点：键值数据库设计、操作  思政元素：开拓创新、勇于探索、积极竞赛，对立积极向上、敬业、诚信、友善的就业风貌，践行社会主义核心价值观的行为准则 | 设计 | 1. 独立完成键值数据库相关内容   2.完成上机报告的撰写并提交 | 目标1  目标2  目标3 |
| 上机 | 文档数据库 | 6 | 重点：MongoDB安装；MongoDB基本操作；  文档数据库的设计；MongoDB案例  难点：MongoDB基本操作；文档数据库设计  思政元素：开拓创新、勇于探索、积极竞赛，对立积极向上、敬业、诚信、友善的就业风貌，践行社会主义核心价值观的行为准则 | 设计 | 1.独立完成文档数据库相关内容 2.完成上机报告的撰写并提交 | 目标1  目标2  目标3 |
| 上机 | 列族数据库 | 6 | 重点：HBase环境搭建；HBase基本操作；  Hbase案例  难点：HBase基本操作  思政元素：开拓创新、勇于探索、积极竞赛，对立积极向上、敬业、诚信、友善的就业风貌，践行社会主义核心价值观的行为准则 | 设计 | 1.独立完成列族数据库内容 2.完成上机报告的撰写并提交 | 目标1目标2 |
|  | 备注： 项目类型填写验证、综合、设计、训练等。 | | | | | |

**五、学生学习成效评估方式及标准**

考核与评价是对课程教学目标中的知识目标、能力目标和素质目标等进行综合评价。在本课程中，学生的最终成绩是由平时成绩、大作业成绩等两个部分组成。

1.平时成绩（占总成绩的40%）：采用百分制。平时成绩分作业（占10%）、上机成绩（占20%）和考勤（占10%）三个部分。评分标准如下表：

|  |  |
| --- | --- |
| **等级** | **评 分 标 准** |
| **1.作业；2.上机；3.考勤** |
| 优秀  （90～100分） | 1.作业逻辑清晰、内容完整；90％以上的习题解答正确或实验习题结果准确无误。  2.上机态度认真，实验报告逻辑清晰、内容完整；能够完成90％以上的上机操作。  3.上课态度认真，积极参与课堂互动；考勤到课率95%以上，不迟到，不早退， 无违纪行为。 |
| 良好  （80～89分） | 1.作业逻辑清晰、内容完整；80％以上的习题解答正确或实验习题结果准确无误。  2.上机态度认真，实验报告逻辑清晰、内容完整；能够完成80％以上的上机操作。  3.上课态度认真，积极参与课堂互动；考勤到课率90%以上，不迟到，不早退， 无违纪行为 |
| 中等  （70～79分） | 1.作业逻辑较清晰、内容较完整；70％以上的习题解答正确或实验习题结果准确无误。  2.上机态度认真，实验报告逻辑较清晰、内容较完整；能够完成70％以上的上机操作。  3.上课态度认真，能够参与课堂互动；考勤到课率85%以上，无违纪行为。 |
| 及格  （60～69分） | 1.作业逻辑不够清晰、内容不太完整；60％以上的习题解答正确或实验习题结果准确无误。  2.上机态度较认真，实验报告逻辑不够清晰、内容不太完整；能够完成60％以上的上机操作。  3.上课态度一般，较少参与课堂互动；考勤到课率80%以上。 |
| 不及格  （60以下） | 1.作业逻辑混乱、内容较少；超过40％的习题解答不正确或实验习题结果错误。  2.上机态度不太认真，实验报告逻辑混乱、内容较少；超过40％的上机操作未完成。  3.上课态度不太认真，较少参与课堂互动；考勤到课率80%以下。 |

2.大作业（占总成绩的60%）：采用百分制度。大作业的考核内容、题型和分值分配情况请见下表：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **考核**  **模块** | **考核内容** | **主要**  **题型** | **支撑目标** | **分值** |
| 大数据存储基础 | 大数据概念；分布式系统；Hadoop分布式处理基础；大数据数据处理基础；大数据存储技术路线 | 简答题；填空题 | 目标1  目标2  目标3 | 15 |
| 数据管理技术发展 | 数据管理系统发展；关系型数据库；NoSQL发展 | 简答题；填空题；程序设计题 | 目标1  目标2  目标3 | 15 |
| NoSQL数据模型 | NoSQL数据库的性质；NoSQL数据库实例介绍 | 简答题 | 目标1  目标2  目标3 | 20 |
| 键值数据库 | 键值数据库的特征；键值数据库的设计与局限 | 简答题、程序设计题 | 目标1  目标2  目标3 | 10 |
| 文档数据库 | 文档数据库数据结构、基本操作；文档数据库设计；MongoDB操作 | 简答题、程序设计题 | 目标1  目标2  目标3 | 20 |
| 列族数据库 | HBase环境搭建；HBase基本操作；Hbase案例 | 简答题、程序设计题 | 目标1  目标2  目标3 | 20 |

|  |  |  |
| --- | --- | --- |
| **序号** | **教学安排事项** | **要 求** |
| 1 | 授课教师 | 职称：初级或者中级及以上  学历（位）：硕士研究生或以上  其他：无 |
| 2 | 课程时间 | 周次：16  节次：3 |
| 3 | 授课地点 | ☑教室 ☑实验室 □室外场地  □其他： |
| 4 | 学生辅导 | 线上方式及时间安排：企业微信，根据学生需要及时辅导  线下地点及时间安排：教室、实验室、办公室，根据学生需要及时辅导 |

1. **教学安排及要求**

**七、选用教材**

[1]柳俊,周苏.大数据存储-从SQL到NoSQL[M].北京: 清华大学出版社,2021年8月.

**八、参考资料**

[1]郭旦怀.大数据存储(NoSQL) [M].北京:清华大学出版社,2023年9月.

[2]周苏.大数据导论(第二版) [M].北京: 清华大学出版社,2022年7月.

**网络资料**

[1]厦门大学数据库实验室网站，https://dblab.xmu.edu.cn/blog/tag/nosql/

大纲执笔人：巩泊成

讨论参与人：杨来

系（教研室）主任：陈强

学院（部）审核人：郭松