**《大数据概论》课程教学大纲**

**一、课程简介**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **课程中文名** | **大数据概论** | | | | | | |
| **课程英文名** | **Introduction to Big Data** | | | | **双语授课** | | □是 ☑否 |
| **课程代码** |  | **课程学分** | **1** | **总学时数** | | 16（含实践0） | |
| **课程类别** | □通识教育课程  □公共基础课程  ☑专业教育课程  □综合实践课程  □教师教育课程 | **课程性质** | □必修  ☑选修  □其他 | **课程形态** | | □线上  ☑线下  □线上线下混合式  □社会实践  □虚拟仿真实验教学 | |
| **考核方式** | □闭卷 □开卷 ☑课程论文 □课程作品 □汇报展示 □报告  ☑课堂表现 □阶段性测试 ☑平时作业 □其他 （可多选） | | | | | | |
| **开课学院** | 大数据与智能工程学院 | | **开课**  **系(教研室)** | 数据科学与大数据技术 | | | |
| **面向专业** | 土木工程专业 | | **开课学期** | 第4学期 | | | |
| **课程负责人** | 程东东 | | **审核人** | 黄金龙 | | | |
| **先修课程** | 无 | | | | | | |
| **后续课程** | 无 | | | | | | |
| **选用教材** | 孟宪伟、许贵秋. 大数据导论[M].人民邮电出版社，2019 | | | | | | |
| **参考书目** | 1．张尧学. 大数据导论[M]. 机械工业出版社, 2018  2．刘鹏. 大数据导论[M]. 清华大学出版社, 2018  3．托马斯.埃尔（著）、彭智勇（译）. 大数据导论[M]. 机械工业出版社, 2017 | | | | | | |
| **课程资源** | 1．中国大数据. 网易公开课（斯坦福大学公开课）  http://open.163.com/movie/2017/6/N/A/MCMM6TB6L\_MCMM75ANA.htm | | | | | | |
| **课程简介** | 大数据概论是该专业的选修课程，包括大数据、云计算和人工智能等概念，通过本课程的学习，使学生全面了解大数据技术与应用的知识体系，掌握大数据思维与算法、数据采集、数据预处理、数据存储、数据挖掘、数据可视化等模块的基本知识。理解大数据技术概念、作用、大数据构架方法，掌握大数据在相关工程问题中的定义、建模、分析和表达。 | | | | | | |

**二、课程目标**

**表 1 课程目标**

|  |  |
| --- | --- |
| **序号** | **具体课程目标** |
| **课程目标 1** | 全面了解大数据技术与应用的知识体系、研究范畴和研究方向，掌握大数据思维与算法、数据采集、数据预处理、数据存储、数据挖掘、数据可视化等模块的基本知识。理解大数据技术概念、作用、大数据架构方法，掌握大数据在相关工程问题中的定义、建模、分析和表达。 |
| **课程目标 2** | 具备使用系统的方法去认知和学习学科专业知识、构建学科体系框架的能力，具备文献检索、资料查询和分析的基本能力，具有土木工程专业相关复杂工程问题与社会可持续发展的客观评价能力和反思能力。 |
| **课程目标** 3 | 开阔学术视野，激发学习本专业的兴趣，树立学习本专业的责任感和自豪感；具备系统分析土木工程专业工程与社会法律法规、可持续发展等问题的计算思维方式；具备良好的文献阅读和拓展分析思维、自主学习和终生意识。 |

**表2-1 课程目标与毕业要求对应关系**

| **毕业要求** | **指标点** | **课程目标** |
| --- | --- | --- |
| **毕业要求7：**环境和可持续发展：能够理解和评价针对土木工程领域复杂工程问题的工程实践对环境、社会可持续发展的影响。【M】 | 7.2 能够站在环境和社会可持续发展的角度思考土木工程实践的可持续性，正确认识和理解土木工程实践可能对环境保护与可持续发展产生不利影响，从而在土木工程实践环节主动融入环保、绿色、节能等可持续发展理念。 | 课程目标2 |
| **毕业要求2：**终身学习：具有自主学习和终身学习的意识，有不断学习和适应行业发展的能力。【H】 | 12.1 能够正确认识自主学习和终身学习的重要性，具有不断学习和适应土木工程专业领域发展的能力。 | 课程目标1  课程目标3 |

**三、课程学习内容与方法**

**（一）理论学习内容及要求**

**表3-1 课程目标、学习内容和教学方法对应关系**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **序号** | **课程模块** | **学习内容** | **学习任务** | **课程目标** | **学习重点难点** | **教学方法** | **学时** |
| 1 | 大数据技术概述 | 1. 大数据技术介绍 | 1.拓展阅读：大数据相关书籍或视频 | 课程目标1 | 重点：  1.大数据技术的定义和研究范畴  难点：  2.大数据技术的研究范畴 | 1.讲授法：能够引导学生对该课程的认知。 | 2 |
| 2. 大数据技术研究范畴 |
| 2 | 大数据的发展和概念 | 1.解读人类信息文明和信息技术发展史 | 1.个人作业：描述自己接触到的大数据应用场景 | 课程目标3 | 重点：  1.人类信息文明和信息技术的发展史  2.大数据时代下，信息技术如何发展  难点：  3.大数据概念和4V特征  4.大数据的应用及价值 | 1.讲授法：能够引导学生的大数据思维  2.专题研讨：能够促进学生了解大数据的应用场景。 | 4 |
| 2.解读大数据时代，讲解大数据的概念和4V特征 | 课程目标1 |
| 3.大数据在社会各领域的应用及体现出的价值 | 课程目标2 |
| 3 | 大数据技术基础 | 1.计算机操作系统及其作用 | 1.个人作业：如何治理大数据 | 课程目标1 | 重点：  1.计算机操作系统的概念及作用  2. Hadoop、Spark平台的发展历程以及两个平台的区别难点：  3.传统数据库和NoSQL及NewSQL的联系和区别  4. 熟知常用的编程语言和大数据算法 | 1.讲授法：能够引导学生理解大数据的相关技术 | 4 |
| 2.计算机编程语言和算法，以及大数据常用的编程语言和大数据算法 | 课程目标2 |
| 3.传统数据库和NoSQL及NewSQL的联系和区别 |
| 4.Hadoop、Spark平台的发展历程以及两个平台的区别 | 课程目标3 |
| 4 | 数据采集与预处理 | 1.大数据的三大数据来源：商业数据、互联网数据和物联网数据 | 1.拓展阅读：阅读数据采集相关书籍。 | 课程目标1 | 重点：  1.商业数据、互联网数据和物联网数据  难点：  2.数据采集方法和原理  3.数据清理、数据集成、数据变换和数据规约 | 1.讲授法：能够引导学生理解数据采集与预处理相关知识和技术 | 2 |
| 2.系统日志方法Flume和网页数据采集方法和原理 |
| 3.数据预处理的四大任务：数据清理、数据集成、数据变换和数据规约 |
| 5 | 数据挖掘 | 1.数据挖掘概念及其对象和价值 | 1.拓展阅读：阅读数据挖掘相关书籍 | 课程目标2 | 重点：  1.数据挖掘概念及其对象和价值  2.数据挖掘工具及其在各领域的应用  难点：  3.数据挖掘算法及其原理和应用场景 | 1.讲授法：能够引导学生了解数据挖掘相关概念和算法 | 2 |
| 2.数据挖掘算法及其原理和应用场景 |
| 3.数据挖掘工具的种类及其在各领域的应用 | 课程目标3 |
| 6 | 数据可视化 | 1.可视化概念定义及作用 | 1.拓展阅读：阅读数据可视化相关书籍 | 课程目标2 | 重点：  1.可视化概念定义及作用  2.大数据可视化案例，及数据可视化面临的发展方向和挑战  难点：  3.可视化技术：科学可视化、信息可视化、可视化分析学，及其应用 | 1.讲授法：能够引导学生了解数据可视化相关概念和技术 | 2 |
| 2.可视化技术：科学可视化、信息可视化、可视化分析学，及其应用场景 |
| 3.大数据可视化案例，及数据可视化面临的发展方向和挑战 | 课程目标3 |

## 四、课程考核

**（一）考核内容与考核方式**

**表4 课程目标、考核内容与考核方式对应关系**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **课程目标** | **考核内容** | **所属**  **学习模块/项目** | **考核占比** | **考核方式** |
| 课程  目标 1 | 1. 大数据技术与应用专业的定义、学科主流的研究方向综合介绍 | 模块1 | 40% | 课程论文 |
| 2.解读大数据时代，讲解大数据的概念和4V特征 | 模块2 |
| 3.计算机操作系统及其作用 | 模块3 |
| 4.大数据的三大数据来源：商业数据、互联网数据和物联网数据；系统日志方法Flume和网页数据采集方法和原理 | 模块4 |
| 课程  目标 2 | 1. 大数据在社会各领域的应用及体现出的价值 | 模块2 | 40% | 课程论文 |
| 2. 计算机编程语言和算法，以及大数据常用的编程语言和大数据算法；传统数据库和NoSQL及NewSQL的联系和区别 | 模块3 |
| 3. 数据挖掘概念及其对象和价值；数据挖掘算法及其原理和应用场景 | 模块5 |
| 4. 可视化概念定义及作用；视化技术：科学可视化、信息可视化、可视化分析学，及其应用场景 | 模块6 |
| 课程  目标 3 | 1. 解读人类信息文明和信息技术发展史 | 模块2 | 20% | 课程论文 |
| 2. Hadoop、Spark平台的发展历程以及两个平台的区别 | 模块3 |
| 3. 数据挖掘工具的种类及其在各领域的应用 | 模块5 |
| 4. 大数据可视化案例，及数据可视化面临的发展方向和挑战 | 模块6 |

**表4-2 课程目标与考核方式矩阵关系**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 课程  目标 | 考核方式 | | | 考核占比 |
| 课程论文成绩比例60% | 课堂表现成绩比例20% | 平时作业成绩比例20% |
| 课程目标1 | 40% | 40% | 40% | 40% |
| 课程目标2 | 40% | 40% | 40% | 20% |
| 课程目标3 | 20% | 20% | 20% | 20% |

## （二）成绩评定

**1.平时成绩评定**

**（1）课堂表现（50%）**：通过学生在课堂上的表现情况、发言与提问情况，来评价学生相关的能力。

**（2）作业完成情况（50%）**：围绕课程的学习目标进行作业的设计。如让学生简述对知识的认识，考核学生对于概念的理解情况，帮助学生将定义转化为自己的理解。

**2.期末成绩评定**

期末的课程论文主要考察学生对大数据技术的知识体系、大数据领域的认识和构架的基本概念等，要求学生了解大数据来源并掌握数据预处理原理及方法，了解大数据计算框架，理解数据挖掘含义，掌握数据挖掘的原理及应用场景，了解数据可视化的定义及其作用和可视化工具，理解大数据与人工智能、云计算之间的的关系，掌握文献检索、资料查询和分析的基本能力，掌握专业相关复杂工程问题与社会可持续发展的客观评价能力和反思能力。考核方式为课程论文。

**3.总成绩评定**

总成绩应由平时考核成绩和期末考核成绩构成。

总成绩（100%）=平时成绩（40%）+期末成绩（60%）。

## （三）评分标准

**表5 评分标准（非试卷考核项目）**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **考核项目** | **评分标准** | | | | |
| **优秀**  **(100>x≥90)** | **良好**  **(90> x≥80)** | **中等**  **(80> x≥70)** | **及格**  **(70> x≥60)** | **不及格**  **(x <60)** |
| 课堂表现 | 课上严格遵守课堂纪律，按要求完成课堂练习且正确率高，积极参与课堂讨论，讨论观点有新意，有见解，不迟到早退，与老师积极互动。 | 课上严格遵守课堂纪律，按要求完成课堂练习且正确率较高，积极参与课堂讨论，讨论观点有一定新意和见解，不迟到早退，与老师积极互动。 | 课上严格遵守课堂纪律，按要求完成课堂练习且有一定正确率，参与课堂讨论，讨论观点有一定新意和见解，不迟到早退，与老师互动。 | 课上严格遵守课堂纪律，按要求完成课堂练习且基本正确率，参与课堂讨论，讨论观点基本的新意和见解，不迟到早退，与老师偶尔互动。 | 课上不遵守课堂纪律，不参与课堂讨论，未按要求完成课堂练习或正确率低。经常迟到早退 |
| 平时作业 | 按照作业目标完成，步骤完整、数据详实、按时提交。 | 按照作业目标完成，步骤基本完整、数据基本详实、按时提交。 | 按照作业目标完成，步骤完整，缺乏数据支撑，按时提交 | 按照作业目标完成，步骤描述不完整，缺乏数据支撑，按时提交 | 未按照作业目标完成，缺乏步骤描述、数据支撑，未按时提交。 |
| 课程论文 | （1）论文选题符合课程性质，选题范围适中，具有较高的研究价值和意义，表现出很强的问题意识。（2）论证过程严谨，所使用的证据或材料充分，结论清晰，具有相当的说服力和解释力。（3）文章结构合理，组织严密，连贯一致。（4）语言表达准确，叙述清楚，所使用的教育专业术语规范。（5）论文符合学术规范。 | （1）论文选题恰当合理，具有较高的研究价值和意义，表现出较强的问题意识。（2）论证过程较为严谨，所使用的证据或材料较为充分，结论清晰，具有较强的说服力和解释力。（3）文章结构合理，组织较为严密，连贯一致。（4）语言表达较为准确，叙述清楚，所使用的教育专业术语较为规范。（5）论文基本符合学术规范，无明显错误。 | （1）论文选题较为合理，具有一定的研究价值和意义，表现出一定的问题意识。（2）论证过程具有一定的严谨性，所使用的证据或材料较为充分，结论清晰，具有一定的说服力和解释力。（3）文章结构较为合理，组织较为严密。（4）语言表达较为准确，叙述较为清楚，所使用的教育专业术语较为规范。  （5）论文基本符合学术规范，有部分错误。 | （1）论文主题具有一定的研究价值和意义，但选题凝练不够，问题意识欠佳。（2）论证过程较为合理但不太严谨，具有一定的证据或材料但不够充分，结论基本清晰。（3）文章结构较为合理，组织具有一定的严密性，但存在部分不连贯现象。（4）语言表达基本清楚，所使用的教育专业术语基本规范。（5）论文基本符合学术规范，有部分错误。 | （1）论文选题不符合课程性质，或主题不明确（2）论证过程随意，所使用的证据或材料极其不充分，结论不清晰。（3）文章结构混乱，存在前后不连贯现象。（4）语言不通顺，所使用的教育专业术语不规范。（5）论文明显不符合学术规范，或存在抄袭现象。 |

## 五、其它说明

本课程大纲依据2023版土木工程专业人才培养方案，由大数据与智能工程学院（部）数据科学与大数据技术教学系（教研室）讨论制定，土木学院（部）教学工作委员会审定，教务处审核批准，自2023级开始执行。