

《大数据基础》课程教学大纲

课程英文名	Big Data Foundations、Base of Big Data				
课程代码	B0505300	课程类别	专业课	课程性质	专业选修
学 分	2		总学时数	32	
开课学院	计算机学院		开课基层教学组织	智能信息处理课程组	
面向专业	计算机科学与技术、计算机科学与技术(第二学士学位)、软件工程、智能财务(软件工程)、智能计算与数据科学(计算机科学与技术)、信息与计算科学		开课学期	第 5 学期	

注：课程类别是指学科基础课/专业课/实践环节/通识公共课/公共基础课/其他；

课程性质是指通识必修/通识选修/学科必修/专业必修/专业选修/实践必修/实践选修。

一、 课程目标

大数据时代的来临使人类拥有更多的机会和条件在各个领域更深入地获得和使用全面数据、完整数据和系统数据，深入探索现实世界的规律。大数据在目前计算机领域中有着极其重要的地位和作用，为学生今后从事计算机相关领域的工作打下较坚实的知识及能力基础。课程主要以大数据系统的基础原理为基础，使学生较全面、深入地理解和掌握大数据系统的基本概念、基本方法、主要功能及实现技术，具备初步的分析、设计和开发大数据系统的能力，培养学生解决大数据问题的能力。同时，结合国家建设和民族复兴的新时代背景，增强学生家国情怀与文化自信，激发学生使命感和责任心。

通过理论教学和实践活动，达到以下课程目标：

课程目标 1：能够运用大数据基本知识，分析 Hadoop 各功能模块的实现原理与技术。

课程目标 2：能够应用现代信息技术工具，通过文献检索及研究，分析大数据开发技术及思路，并能针对大数据问题提出初步的解决建议。

课程目标 3：掌握大数据的新技术与发展趋势，初步具备分析并设计大数据系统架构及功能目标能力。

课程目标 4：具备基本的科学素养，及时了解国内外大数据新技术的发展趋势，及时掌握国家相关方面的科技战略需求，树立强烈的爱国主义使命感与责任心。

二、 课程目标与毕业要求对应关系

课程的课程目标对计算机科学与技术专业毕业要求指标点的支撑情况如表 1 所示：

表 1 课程目标与毕业要求对应关系

毕业要求	指标点	课程目标及支撑权重
毕业要求 1：工程知识：掌握数学、自然科学、工程基础、计算机专业领域的知识，并能应用于计算机领域复杂工程问题的解决方案中。	1-1 掌握数学、自然科学和工程基础知识，并能够用于计算机领域复杂工程问题的理解、描述、推理与模型选择。	目标 2：0.5 目标 3：0.5
	1-2 掌握计算机科学核心知识与理论，能够针对计算机领域复杂工程问题建立模型，并利用模型解决问题。	目标 1：0.5 目标 3：0.5
	1-4 掌握计算机系统、物联网、人工智能、大数据、网络安全等某个专业领域的知识，并用于解决计算机领域的复杂工程问题。	目标 1：0.5 目标 3：0.5
毕业要求 2：问题分析：能够应用数学、自然科学和工程科学的基本原理，以及科学思维方法，对计算机领域的复杂工程问题进行识别、表达和分析，并通过文献查阅与研究获得有效结论。	2-2 具备文献检索能力，能够对计算机领域的复杂工程问题进行调查、研究分析。	目标 1：0.3 目标 2：0.3 目标 3：0.4
毕业要求 5：使用现代工具：能够针对计算机领域的复杂工程问题，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，包括对计算机领域的复杂工程问题进行预测与模拟，并能够理解其局限性。	5-1 了解计算机领域常用的现代工程工具和信息技术工具的适用范围、使用原理与方法，理解其局限性。	目标 1：0.5 目标 2：0.5
毕业要求 12：终身学习：具有自主学习和终身学习的意识，有不断学习和适应信息技术高速发展的能力。	12-2 具备自主学习的能力，包括技术理解力、归纳总结能力和提出问题的能力。	目标 1：0.2 目标 2：0.3 目标 3：0.3 目标 4：0.2

课程的课程目标对软件工程相关专业毕业要求指标点的支撑情况如表 2 所示：

表 2 课程目标与毕业要求对应关系

毕业要求	指标点	课程目标及支撑权重
毕业要求 1：工程知识：掌握数学、自然科学、工程基础、软件工程专业领域的知识，并能应用于软件工程领域复杂工程问题的解决方案中。	1-1 掌握数学、自然科学和工程基础知识，并能够用于软件工程领域复杂工程问题的理解、描述、推理与模型选择。	目标 2：0.5 目标 3：0.5
	1-2 掌握软件工程核心知识与理论，能够针对软件工程领域复杂工程问题建立模型，并利用模型解决问题。	目标 1：0.5 目标 3：0.5
	1-4 掌握某个专业领域知识，并用于解决软件工程领域复杂工程	目标 1：0.5 目标 3：0.5

	问题。	
毕业要求 2: 问题分析: 能够应用数学、自然科学、工程科学的基本原理和科学思维方法, 对软件工程领域复杂工程问题进行识别、表达和分析, 并通过文献查阅与研究获得有效结论。	2-2 具备文献检索能力, 能够对软件工程领域复杂工程问题进行调查、研究分析。	目标 1: 0.3 目标 2: 0.3 目标 3: 0.4
毕业要求 5: 使用现代工具: 能够针对软件工程领域复杂工程问题, 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, 包括对软件工程领域复杂工程问题进行预测与模拟, 并能够理解其局限性。	5-1 了解软件工程领域常用的现代工程工具和信息技术工具的适用范围、使用原理与方法, 理解其局限性。	目标 1: 0.5 目标 2: 0.5
毕业要求 12: 终身学习: 具有自主学习和终身学习的意识, 有不断学习和适应信息技术高速发展的能力。	12-2 具备自主学习的能力, 包括技术理解力、归纳总结能力和提出问题的能力。	目标 1: 0.2 目标 2: 0.3 目标 3: 0.3 目标 4: 0.2

课程的课程目标对智能财务(软件工程)专业毕业要求指标点的支撑情况如表 3 所示:

表 3 课程目标与毕业要求对应关系

毕业要求	指标点	课程目标及支撑权重
毕业要求 1: 工程知识: 掌握数学、自然科学、工程基础、财务领域基础知识和软件工程专业领域知识, 并能应用在智能财务软件领域复杂工程问题的解决方案中。	1-1 掌握数学、自然科学和工程基础知识, 并能够用于智能财务软件工程领域复杂工程问题的理解、描述、推理与模型选择。 1-2 掌握软件工程和财务核心知识与理论, 能够针对智能财务软件工程领域复杂工程问题建立模型, 并利用模型解决问题。 1-4 掌握某个专业领域知识, 并用于解决智能财务软件工程领域复杂工程问题。	目标 2: 0.5 目标 3: 0.5 目标 1: 0.5 目标 3: 0.5 目标 1: 0.5 目标 3: 0.5
毕业要求 2: 问题分析: 能够应用数学、自然科学、财务领域基础知识和工程科学的基本原理, 以及计算科学思维方法, 对智能财务软件领域复杂工程问题进行识别、表达和分析, 并通过文献查阅与研究获得有效结论。	2-2 具备文献检索能力, 能够对智能财务软件工程领域复杂工程问题进行调查、研究分析。	目标 1: 0.3 目标 2: 0.3 目标 3: 0.4
毕业要求 5: 使用现代工具: 能够针对智能财务软件领域复杂工程问题, 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, 包括对复杂工程问题的预测与模拟, 并能够理解其局限性。	5-1 了解智能财务软件工程领域常用的现代工程工具和信息技术工具的适用范围、使用原理与方法, 理解其局限性。	目标 1: 0.5 目标 2: 0.5

毕业要求 12: 终身学习: 具有自主学习和终身学习的意识, 有不断学习和适应信息技术高速发展的能力。	12-2 具备自主学习的能力, 包括技术理解力、归纳总结能力和提出问题的能力。	目标 1: 0.2 目标 2: 0.3 目标 3: 0.3 目标 4: 0.2
---	---	--

课程的课程目标对智能计算与数据科学(计算机科学与技术)专业毕业要求指标点的支撑情况如表 3 所示:

表 4 课程目标与毕业要求对应关系

毕业要求	指标点	课程目标及支撑权重
毕业要求 1: 工程知识: 能够掌握数学、自然科学、工程基础、计算机软硬件知识、人工智能、智能计算和数据科学的基础理论及专业知识, 并应用在人工智能、智能计算和大数据专业领域及其他相关交叉领域的复杂工程问题的解决方案中。	1-1 掌握数学、自然科学和工程基础知识, 并能够用于计算机领域复杂工程问题的理解、描述、推理与模型选择。	目标 2: 0.5 目标 3: 0.5
	1-2 掌握计算机科学核心知识与理论, 能够针对计算机领域复杂工程问题建立模型, 并利用模型解决问题。	目标 1: 0.5 目标 3: 0.5
	1-4 掌握人工智能、智能计算、大数据等某个专业领域的知识, 并用于解决计算机领域的复杂工程问题。	目标 1: 0.5 目标 3: 0.5
毕业要求 2: 问题分析: 能够应用数学、自然科学、工程科学、人工智能、智能计算和大数据的基本原理, 对人工智能和大数据专业领域及其他相关交叉领域的复杂工程问题进行识别、表达、分析和抽象建模, 并通过文献查阅与研究获得有效结论。	2-2 具备文献检索能力, 能够对计算机领域的复杂工程问题进行调查、研究分析。	目标 1: 0.3 目标 2: 0.3 目标 3: 0.4
毕业要求 5: 使用现代工具: 能够针对复杂工程问题, 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, 包括对复杂工程问题的预测与模拟, 并能够理解其局限性。	5-1 了解计算机领域常用的现代工程工具和信息技术工具的适用范围、使用原理与方法, 理解其局限性	目标 1: 0.5 目标 2: 0.5
毕业要求 12: 终身学习: 具有自主学习和终身学习的意识, 有不断学习和适应信息技术高速发展的能力。	12-2 具备自主学习的能力, 包括技术理解力、归纳总结能力和提出问题的能力。	目标 1: 0.2 目标 2: 0.3 目标 3: 0.3 目标 4: 0.2

三、 课程目标与教学内容和方法的对应关系

《大数据基础》课程目标与教学内容、建议教学方法的对应关系如表 3 所示。教师可根据实际情况和自身特长采用合适的教学方法。

表 5 课程目标与教学内容、教学方法的对应关系

教学内容	教学方法	课程目标
10. 大数据概论	课堂讲授、课堂小讨论、课堂测试	1、2、4
11. 大数据处理架构 hadoop	课堂讲授、课堂小讨论、课堂测试、PBL 教学法、案例分析、视频学习	1、2、4
12. 分布式文件系统 HDFS	课堂讲授、课堂小讨论、课堂测试、案例分析、视频学习	1、2、4
13. 分布式数据库 HBase	课堂讲授、课堂小讨论、课堂测试、案例分析、视频学习	1、2、4
14. NoSQL 数据库	课堂讲授、课堂小讨论、PBL 教学法、案例分析、视频学习	1、2、3、4
15. 云数据库	课堂讲授、课堂小讨论、视频学习	1、2、3、4
16. MapReduce	课堂讲授、课堂小讨论、课堂测试、案例分析、视频学习	1、2、3、4
17. Hadoop 架构再探讨	课堂讲授、课堂小讨论、课堂测试、案例分析、视频学习	1、2、3、4

课程教学的详细内容与要求如下：

1. 大数据概论

(1) 教学内容：

- 大数据时代背景；
- 大数据发展前景；
- 大数据的基本概念：定义、结构类型及核心特征；
- 大数据系统设计目标和原则；
- 大数据处理的关键技术；
- 大数据对企业的挑战性；
- 企业大数据的发展方向；

(2) 教学重点：大数据的基本概念、大数据系统设计目标和原则。

(3) 教学难点：大数据的结构类型和核心特征。

(4) 教学要求：了解大数据的概念、特征以及大数据的结构类型。了解大数据的设计思想、设计目标和设计原则。

思政融合点 1: 引导学生查阅资料，分析并总结国内外大数据技术的发展趋势和特点。通过对比国内外大数据先进技术，激发学生爱国主义使命感、责任心和学习先进技术的热情。

2. 大数据处理架构 Hadoop

(1) 教学内容:

- Hadoop 的概况、优势、发展史及应用前景；
- Hadoop 的功能和作用；
- Hadoop 的体系结构；
- Hadoop 与分布式开发；
- Hadoop 与关系型数据库管理系统；
- Hadoop 应用案例；

(2) 教学重点: Hadoop 的功能和作用；Hadoop 的体系结构的两大核心模块；Hadoop 与分布式开发的实现方式。

(3) 教学难点: Hadoop 体系结构的两大核心模块；Hadoop 与分布式开发的实现方式。**(4) 教学要求:**

能够配置大数据 java 开发的环境，配置 Hadoop 环境，搭建 Hadoop 集群。了解 Hadoop 的功能和作用。了解 Hadoop 体系结构的两大核心模块和 Hadoop 分布式开发的实现方式。

思政融合点 2: Hadoop 社区中有很多来自不同国家的不同公司贡献的组件，调研我国公司如华为、阿里等在 Hadoop 社区中贡献的内容，让学生了解我国大数据相关先进技术面向世界做出的贡献，以激发学生的民族自豪感。

3. 分布式文件系统 HDFS

(1) 教学内容:

- HDFS 简介；
- HDFS 的特性和设计目标；
- HDFS 的核心设计；
- HDFS 的体系结构；
- HDFS 中数据流的读写；
- HDFS 的 HA 机制；
- HDFS 的 Federation 机制；
- HDFS 的命令行常用接口；
- HDFS 的 Java 接口；

(2) 教学重点: HDFS 的设计目标和基本概念；HDFS 的体系结构；HDFS 的运行机制；HDFS 文件系统的命令操作；HDFS 的 Java 接口。

(3) 教学难点: HDFS 的体系结构；HDFS 的运行机制；HDFS 文件系统的命令操作 HDFS

的 Java API。

- (4) **教学要求:** 了解 HDFS 的特征和设计目标。了解 HDFS 中数据流读写的运行机制。能够利用命令行进行 HDFS 文件系统的基本操作。掌握 Java 对 HDFS 的程序开发操作，包含目录管理、文件列表、读取、导入导出数据等开发操作。

4. 分布式数据库 HBase

- (1) **教学内容:**

- HBase 概述；
- HBase 访问接口；
- HBase 数据模型；
- HBase 的实现原理；
- HBase 运行机制；
- HBase 应用方案；
- HBase 编程实践。

- (2) **教学重点:** HBase 数据模型；HBase 的实现原理；HBase 编程实践。

- (3) **教学难点:** HBase 面向列的存储与传统面向行的存储的比较；HBase 的系统架构及各个组成部分的功能、工作原理；HBase 的安装、配置、常见命令及 Java API。

- (4) **教学要求:** 能够理解 HBase 面向列的存储与面向行的存储之间的区别和各自优缺点。能够了解 HBase 的系统架构及各个组成部分，特别是 Region 服务器、Store、HLog 的工作原理。学习 HBase 编程实践，完成 HBase 的安装与配置，了解 HBase 的常见 Shell 命令及 Java API 的使用。

5. NoSQL 数据库

- (1) **教学内容:**

- NoSQL 简介；
- NoSQL 兴起的原因
- NoSQL 与关系数据库的比较；
- NoSQL 的四大类型；
- NoSQL 的三大基石；
- NoSQL 到 NewSQL 数据库；
- 文档数据库 MongoDB

- (2) **教学重点:** NoSQL 与关系数据库的比较；NoSQL 的四大类型；NoSQL 的三大理论基石。

- (3) **教学难点:** NoSQL 数据库和关系型数据库的比较及应用场景；理清 NoSQL 数据库的三大理论基石与关系型数据库理论基石。

(4) 教学要求: 能够理解 NoSQL 数据库和关系型数据库各自的优缺点及应用场景。能够了解键值数据库、列族数据库、文档数据库、图数据库各自的数据模型、典型应用、优缺点及代表产品。能够学习 NoSQL 数据库的三大理论基石，CAP、BASE 和最终一致性。

6. 云数据库

(1) 教学内容:

- 云数据库概述；
- 云数据库产品；
- 云数据库系统架构；
- Amazon AWS 和云数据库；
- 微软云数据库；
- 云数据库实践。

(2) 教学重点: 云数据库系统架构；各类云数据库产品；云数据库实践。

(3) 教学难点: 云数据库系统架构及其中的角色和组件；云数据库实践及本地数据库到云端数据库的迁移。

(4) 教学要求: 了解云数据库的概念；学习云数据库产品的架构及组件；了解各类云数据库产品的结构及使用；学习云数据库实践及本地数据库到云端数据库的迁移。

思政融合点 3: 引导学生了解我国百度、阿里、华为等公司在云数据库开发应用上取得的成就，如百度云、阿里云、华为云等在国内的广泛应用，总结其技术和方法，树立民族自豪感；同时通过介绍亚马逊、谷歌等公司的云数据库产品，引导学生瞄准世界科技前沿，树立投身科学的研究和技术创新的远大理想。

7. MapReduce

(1) 教学内容:

- MapReduce 编程模型简介；
- Hadoop MapReduce 架构；
- MapReduce 程序的设计方法；
- WordCount 编程实例；
- MapReduce 的编程；
- MapReduce 在集群上的运作；
- MapReduce 类型与格式；
- MapReduce 的 Java API 解析；

(2) 教学重点: MapReduce 的设计思想、基本概念；MapReduce 的系统架构、作业运行机制和关键技术；MapReduce 的程序设计方法，能够编写 Mapper 类和 Reducer 类；

MapReduce 在集群上的运作；MapReduce 的数据类型及使用。MapReduce 的 Java API 解析。

- (3) **教学难点：**MapReduce 的设计思想、系统架构；MapReduce 的程序设计方法，能够编写 Mapper 类和 Reducer 类；MapReduce 在集群上的运作；MapReduce 的 Java API 解析。
- (4) **教学要求：**了解 MapReduce 的设计思想与系统架构。通过基础知识的学习，能够编写 MapReduce 的 Mapper 类和 Reducer 类。能够在集群下运作 MapReduce 项目。能够将 HDFS 文件系统中的文件作为输入开发 MapReduce 项目。

思政融合点 4：引导学生了解 MapReduce 并行计算框架的历史，鼓励学生通过现代化手段，以查阅文献、自学、小组研讨等方式，分析并总结 MapReduce 在国内各行业中的实际应用。通过了解并追踪其应用，帮助学生理解科技第一生产力对于国家建设的重要意义。

8. Hadoop 架构再探讨

- (1) **教学内容：**

- Hadoop 的优化与发展；
- YARN 的架构；
- 剖析 MapReduce 作业运行机制；
- 剖析基于 YARN 的运行机制；
- Shuffle 和排序；
- 任务的执行、作业的调度；
- 在 YARN 上运行 MapReduce 实例；

- (2) **教学重点：**YARN 的架构；剖析经典的 MapReduce 作业运行机制；剖析基于 YARN 的 MapReduce 运行机制；基于 YARN 的 MapReduce 的优势；Map 端和 Reduce 端的 shuffle 过程。
- (3) **教学难点：**基于 YARN 的 MapReduce 运行机制。Map 端和 Reduce 端的 shuffle 过程。
- (4) **教学要求：**了解 YARN 的架构。了解基于 YARN 的 MapReduce 的运行机制。使用现代信息技术工具查阅、分析及研究基于 YARN 的 MapReduce 相比于传统 MapReduce 的优势。掌握 shuffle 在 map 端和 reduce 端的过程。

思政融合点 5：总结大数据相关技术的理论及实践，帮助学生理解先进技术在我国各个领域中的具体应用，增强学生家国情怀与文化自信，激发学生使命感和责任心。

四、实践环节及要求

考虑到有配套课程《大数据开发实践》，本课程课内不再安排上机实践。

五、与其它课程的联系

先修课程：程序设计基础、Java 面向对象程序设计。

后续课程：大数据开发实践、大数据应用案例及分析。

六、学时分配

表 6 学时分配表

教学内容	讲课时数	实验时数	实践学时	课内上机时数	课外上机时数	自学时数	习题课	讨论时数
1. 大数据概论	4					2		
2. 大数据处理架构 hadoop	4					2		
3. 分布式文件系统 HDFS	4					4		
4. 分布式数据库 HBase	4					4		
5. NoSQL 数据库	4					4		
6. 云数据库	4					2		
7. MapReduce	4					4		
8. Hadoop 架构再探讨	4					2		
合计	32					16	8	
总计						32 课时		

七、课程目标达成途径及学生成绩评定方法

1. 课程目标达成途径

表 7 课程目标与达成途径

课程目标	达成途径
课程目标 1： 能够运用大数据基本知识，分析 Hadoop 各功能模块的实现原理与技术。	以引导式、启发式和总结式教学方法为主，通过重点/难点内容讲解、课后作业、进行随堂提问、课堂程序演示等模式，帮助学生能够运用大数据基本知识，分析 Hadoop 各功能模块的实现原理与技术。
课程目标 2： 能够应用现代信息技术工具，通过文献检索及研究，分析大数据开发技术及思路，并能针对大数据问题提出初步的解决建议。	以启发式、分析式和研讨式教学方法为主，结合典型应用案例，通过上机实践、作业、随堂提问、课堂讨论等模式，帮助学生具备运用大数据的思维方式，为实际应用问题建立问题模型，具备初步分析、设计和开发大数据系统的初步能力。
课程目标 3： 掌握大数据的新技术与发展趋势，初步具备分析并设计大数据系统架构及功能目标能力。	以引导式、启发式和总结式教学方法为主，通过重点/难点内容讲解、上机实践、课后作业、进行随堂提问、课堂讨论等模式，帮助学生掌握大数据的新技术与发展趋势，初步具备分析并设计大数据系统架构及功能目标能力。
课程目标 4： 具备基本的科学素养，及时了解国内外大数据新技术的发展趋势，及	通过课堂讲授、课后自学、文献查阅、课堂讨论、分析对比、总结报告等各种方式，让学生对当代大数据技术的现状与发展趋势有所了解，建立终生学习的意识；同时，进一步了解目前国内相关先进大数

时掌握国家相关方面的科技战略需求，树立强烈的爱国主义使命感与责任心。	据技术与取得的成就，从而建立强烈的民族自豪感与爱国主义使命感。
------------------------------------	---------------------------------

2.学生成绩评定方法

该课程为考试课程，考试方式为闭卷。该课程采用形成性评价与终结性评价相结合的评价方法，学期总评成绩由二部分构成：平时成绩，占比 50%，任课教师可调整平时成绩组成比例；期末考试成绩，占比 50%。各部分的具体评价环节、关联课程目标、评价依据及方法和在总成绩中的占比，如表 8 所示。

表 8 课程考核与成绩评定方法

成绩构成	考核项目	考核关联的课程目标	考核依据与方法	占总评成绩的比重
平时成绩	课程思政实践	4	基于大数据主题，通过课外文献查阅、课堂展示、课堂小组讨论、阅读报告等多种形式，考查学生对我国相关先进技术的了解情况以及核心价值观状况	5%
	课堂成绩	1、2、3	考勤、课堂讨论、问答等。	10%
	课后作业	1、2、3	作业是否及时完成、作业完成质量。	15%
	实验操作及实验报告	1、2、3、4	布置大数据基础相关的若干个课下实验，包括HDFS、MapReduce 等，提交实验报告。	20%
期末考试	闭卷考试	1、2、3、4	考试成绩	50%
总评成绩		1、2、3、4	=平时成绩*50%+考试成绩*50%	100%

表 9. 考核内容详细评分标准

考核内容	评分标准			
	90-100	75-90	60-75	<60
课程思政实践	报告条理清晰，文字流畅，字数 \geq 4000，参考文献数量 \geq 8 且相关性强；内容完整且材料丰富，体现强烈的使命感、责任心与民族自豪感	报告条理清楚，字数 \geq 3000，参考文献数量 \geq 5 且相关性较好；内容完整，材料不够丰富，能体现学生的使命感、责任心与民族自豪感	报告有一定条理，字数 \geq 1000，参考文献数量 \geq 2 且基本相关；内容基本完整但材料较少，能体现学生的使命感与民族自豪感	报告字数 $<$ 1000，参考文献数量 $<$ 2；内容少，或有抄袭现象，体现不出学生的使命感与民族自豪感
课堂成绩	雨课堂测试、课堂练	雨课堂测试、课堂练	雨课堂测试、课堂练	雨课堂测试、课堂练

	习、回答问题等据实评价；或参与回答次数在教学班前 15%	习、回答问题等据实评价；或参与回答次数在教学班前 50%	习、回答问题等据实评价；或参与回答次数在教学班前 85%	习、回答问题等据实评价；或参与回答次数在教学班后 15%
课后作业	非标作业：方案等设计合理，分析准确，能满足问题全部要求	非标讨论题：方案较合理，分析较正确，能基本满足问题全部要求	非标讨论题：方案基本合理，能满足问题大部分要求	非标讨论题：方案不够合理，只能满足问题少量要求
标准题目：按照作业题目评分标准据实评价				
实验操作及实验报告	程序运行流畅，功能完善，性能好；代码独创性好；算法与数据结构设计或代码实现等有 2 处及以上创新点	程序运行正常，功能基本实现，但不够完善、优化等，代码自创率应高于 50%；算法与数据结构设计或代码实现等至少有一处创新点	程序运行基本正常，有少量 bug；功能实现至少达到 80%，代码自创率应高于 30%；基本没有创新性	程序运行 bug 多，功能实现低于要求的 80%
期末闭卷考试	按照期末试卷评分标准据实评价			

八、 教学资源

表 10 课程的基本教学资源

资源类型	资源
教材	林子雨著，《大数据技术原理与应用》，（第 2 版）人民邮电出版社，2017.07。
参考书籍或文献	(1) 中科普开著，《大数据技术基础》（第 1 版），清华大学出版社，2018.08。 (2) Tom White 等著，《Hadoop 权威指南》（第 3 版 翻译版），清华大学出版社，2015.07。 (3) 陆价恒等著，《Hadoop 实战》（第 2 版），机械工业出版社，2014.12。 (4) 俞东进等，《大数据基础、技术与应用》，科学出版社，2022.02。
教学文档	(1) Hadoop 官方网站 1: http://hadoop.apache.org/

九、 课程目标、毕业要求指标点达成度定量评价

在课程结束后，需要对每一个课程目标（含思政课程目标）进行达成度的定量评价，用以实现课程的持续改进。

课程目标达成度的定量评价算法：

- 1、使用教学活动（如课程思政实践、课后作业、课堂练习、课堂讨论、互动、大作业等等）成绩或期末考试部分题目得分率作为评价项目，来对某个课程目标进行达成度的定量评价；
- 2、为保证考核的全面性和可靠性，要求对每一个课程目标的评价项目选择超过两种；
- 3、根据施教情况，评价项目可以由教师自行扩展，权重比例可以由教师自行设计；
- 4、对某一个课程目标有支撑的各评价项目权重之和为 1；
- 5、使用所有学生（含不及格）的平均成绩计算。

表 11 课程目标达成度定量评价方法

课程目标	课程目标达成度评价方式
课程目标 1: 能够运用大数据基本知识，分析 Hadoop 各功能模块的实现原理与技术。	课堂成绩: 0.3 实验操作及实验报告: 0.2 期末闭卷考试: 0.3 课后作业: 0.2
课程目标 2: 能够应用现代信息技术工具，通过文献检索及研究，分析大数据开发技术及思路，并能针对大数据问题提出初步的解决建议。	课堂成绩: 0.5 实验操作及实验报告: 0.2 期末闭卷考试: 0.3
课程目标 3: 掌握大数据的新技术与发展趋势，初步具备分析并设计大数据系统架构及功能目标能力。	课堂成绩: 0.2 实验操作及实验报告: 0.3 期末闭卷考试: 0.3 课后作业: 0.2
课程目标 4: 具备基本的科学素养，及时了解国内外大数据新技术的发展趋势，及时掌握国家相关方面的科技战略需求，树立强烈的爱国主义使命感与责任心。	实验操作及实验报告: 0.2 期末闭卷考试: 0.5 课程思政实践: 0.3

十、说明

本课程大纲主要用于规范软件工程（卓越工程师计划）、软件工程、软件工程（国际教育学院）以及信息与计算科学专业的《大数据基础》课程的教学目标、教学内容、教学方法、教学要求以及考核评价方法等，承担该课程的教师必须遵照本大纲安排授课计划、实施教学过程，完成学生各个阶段与各方面的学习成果考核与评价；在学期末，需对课程目标和课程支撑的毕业要求指标点进行达成度评价。

本课程大纲于 2021 级开始执行，生效之日原先版本均不再使用。

十一、编制与审核

表 12 大纲编制与审核信息

工作内容	责任部门或机构	负责人	完成时间
编制	智能信息处理组	俞东进、陈洁、孙笑 笑	2022.03
审核	智能信息处理组	彭勇	2022.03
审定	计算机学院教学工作委员会	黄孝喜	2022.05