

大数据工程技术专业人才培养方案

一、专业名称及代码

名称：大数据工程技术

代码：310205

二、入学要求 普通高中毕业生、中等职业学校毕业生或同等学力者

三、学制与学位

学制：标准学制四年

学位：授予工学学士学位

四、职业面向

所属专业大类（代码）	所属专业类（代码）	对应行业	主要职业类别	主要岗位群或技术领域举例	职业技能等级证书举例
电子与信息类 (31)	计算机类 (3102)	大数据应用行业	大数据分析、 大数据系统 架构	大数据应用 开发工程师、大数据 分析师、 大数据系统 架构师	大数据应用 开发（Java） 证书

五、人才培养目标与培养规格

（一）培养目标

大数据工程技术专业培养目标：本专业培养德、智、体、美、劳全面发展，践行社会主义核心价值观，具有良好的职业道德和人文素养，具备大数据平台架构设计与运行维护、数据建模和分析以及解决行业应用问题的能力，信息化时代的终身学习能力，能够承担大数据平台架构设计、数据采集、存储与管理、数据分析与可视化任务，能在相应领域从事各行业大数据分析、处理、服务、开发和利用等工作，具有社会责任感、创新精神和较强实践能力，服务地方经济的大数据应用行业的高层次技术技能型人才。

（二）培养规格

1. 素质

(1) 具有正确的世界观、人生观、价值观。坚决拥护中国共产党领导，树立中国特色社会主义共同理想，践行社会主义核心价值观，具有深厚的爱国情感、国家认同感、中华民族自豪感；崇尚宪法、遵守法律、遵规守纪；具有社会责任感 and 参与意识。

(2) 具有良好的职业道德和职业素养。崇德向善、诚实守信、爱岗敬业，具有精益

求精的工匠精神；尊重劳动、热爱劳动，具有较强的实践能力；具有质量意识、绿色环保意识、安全意识、信息素养、创新精神；具有较强的集体意识和团队合作精神，能够进行有效的人际沟通和协作，与社会、自然和谐共处；具有职业生涯规划意识。

(3) 具有良好的身心素质和人文素养。具有健康的体魄和心理、健全的人格，能够掌握基本运动知识和一两项运动技能；掌握一定的学习方法，具有良好的生活习惯、行为习惯和自我管理能力。

2. 知识

(1) 公共基础知识

掌握必备的思想政治理论、科学文化基础知识、中华优秀传统文化知识。

(2) 专业知识

本专业要求具有从事数据科学与大数据技术行业所需的数学、自然科学、工程基础和专业基础知识，并能够将这些知识用于解决大数据应用领域复杂工程问题。掌握大数据系统设计，数据平台的搭建、维护和优化，为设计复杂大数据工程问题的解决方案提供支持。

3. 能力

- (1) 具有探究学习、终身学习、分析问题、解决问题的能力。
- (2) 具有良好的语言、文字表达能力、沟通能力。
- (3) 具备本专业必需的信息技术应用和维护能力。
- (4) 运用自然科学与人文社科知识，完成对复杂大数据应用系统进行综合分析。
- (5) 能够在复杂大数据工程问题的设计/开发环节中体现创新意识。
- (6) 根据大数据工程问题的特征和研究目标选择研究路线，设计测试方案。
- (7) 能够设计大数据领域复杂工程问题的解决方案，并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。
- (8) 能够理解和评价针对数据科学与大数据技术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。
- (9) 具有人文社会科学素养、社会责任感，能够在数据科学与大数据技术工程实践中理解并遵守工程职业道德和规范，履行责任。

六、职业技能等级证书

大数据应用开发（Java）证书

七、毕业要求

修完教学计划规定的全部课程和实践环节，取得 174.5 学分；达到国家大学生体质健康标准，准予毕业。

符合《中华人民共和国学位授予条例》和我校学位授予相关规定，可授予学位。

八、专业核心课程简介

课程名称	面向对象程序设计 (Java)	学分	4
教学目标	本课程要求学生能够了解 Java 语言的基本使用规则及语法规则,具备 Java 程序设计语言相关知识、能够编写、调试 Java 程序,遵守良好的代码编写规范,能进行简单的 Java 应用程序开发,为后续课程打下坚实基础。		
教学内容	Java 开发环境搭建、java 语言基础、变量定义使用、流程控制、简单算法、Java 面向对象基础、封装、继承、多态、接口、字符串操作、常用类,集合与容器、GUI 编程基础等内容。		
实训项目	<ol style="list-style-type: none"> 1. 三个程序结构的实训。 2. 类的定义与对象的创建。 3. Java 面向对象编程测试。 4. JavaGUI 开发基础。 5. 企业提供的 Javawing 项目实训。 6. 通过 java 基础知识的学习能够独立开发控制台 DVD 管理系统。 		
课程名称	大数据技术架构	学分	3
教学目标	本课程要求学生了解大数据的基本概念,理解大数据处理需求,学会大数据的云中计算,掌握大数据计算框架和大数据系统架构。		
教学内容	大数据概念、大数据处理需求、大数据的云中计算、大数据计算框架、大数据系统架构等内容		
实训项目	<ol style="list-style-type: none"> 1. 企业提供的 2. 企业提供的 3. 企业提供的 4. 企业提供的 		
课程名称	数据挖掘	学分	3
教学目标	本课程要求学生了解数据挖掘的重要性与国内外的状况及未来发展方向;掌握数据挖掘的一些基本概念和基本方法,学习和掌握常用的数据挖掘方法,能够熟练地运用数据挖掘技术及工具进行具体数据的挖掘分析。		
教学内容	数据挖掘的基本概念和功能、数据预处理基本知识和方法、数据仓库的设计、分类与预测基本知识和方法、聚类分析的基本概念和方法等。		
实训项目	<ol style="list-style-type: none"> 1、企业提供的 2、企业提供的 		

	3、 企业提供的 基于 sklearn 数据分类挖掘实训 4、 企业提供的 基于 sklearn 数据聚类挖掘实训 5、 企业提供的 在 Python 程序中引用 sklearn 实训		
课程名称	分布式计算框架	学分	3
教学目标	本课程要求学生掌握分布式计算框架的基础知识,理解 MapRuduce、Spark、Storm 的基本知识和相关技术,能够利用 MapRuduce、Spark、Storm 技术实现分布式计算框架的选型。		
教学内容	数据计算演化、分布式计算框架概述、MapRuduce、Spark、Storm、分布式计算框架选型。		
实训项目	1、 企业提供的 MapRuduce 实训 2、 企业提供的 Spark 实训 3、 企业提供的 Storm 实训 4、 企业提供的 分布式计算框架的选型实训		
课程名称	数据可视化	学分	2
教学目标	本课程要求学生了解大数据可视化的工作原理和使用方法,使学生具有大数据可视化的设计和可视化开发的能力,具备大数据可视化编码的基本技能,为将来从事大数据相关领域的工作打下坚实的基础。		
教学内容	大数据可视化的概述、数据可视化的作用、数据可视化设计原则与技巧、可视化图表、文本和网络可视化以及空间信息可视化、可视化工具等。		
实训项目	1、Excel 可视化应用 2、 完全由企业工程师负责完成 Echarts 可视化应用 3、 完全由企业工程师负责完成 Tableau 可视化应用 4、 完全由企业工程师负责完成 pyecharts 可视化应用 5、 完全由企业工程师负责完成 Python 可视化库的应用		
课程名称	应用统计与试验设计	学分	3
教学目标	学生能够理解数理统计的基本概念与理论,掌握统计数据的基本方法,并通过 R 语言实现,为后续的学习、研究和发展奠定必要的基础;培养学生工程实践能力,能够根据需求,分析、处理数据,并设计、实现 R 程序,实现数据分析与编程的全过程;提升学生的专业素质,养成良好的学习态度和学习习惯,提升分析问题解决问题的能力。		

教学内容	统计软件与概率计、数理统计初步与模拟计算、参数估计、检验假设		
实训项目	1. 完全由企业工程师负责完成 基于 R 软件的消费水平统计分析 2. 完全由企业工程师负责完成 基于 R 软件的农户收入水平参数估计 3. 完全由企业工程师负责完成 关于身高寿命假设检验的统计分析		
课程名称	分布式实时计算	学分	3
教学目标	理解分布式实时计算的基本概念与理论，并将其合理应用到实际情况中，在遇到问题的时候采用小组讨论或文献查找等方式，解决问题；总结分布式实时计算的思想，理解实时计算中的基本概念、应用场景，并能基于现阶段最主流的大数据技术如 Kafka、Flink 等解决实际的问题，具备一定的实际项目开发能力；培养学生工程实践能力，能够根据需求，分布式实时计算模型设计与搭建的全过程；提升学生的专业素质，养成良好的学习习惯，提升分析问题解决问题的能力。		
教学内容	本课程将系统讲授分布式实时计算原理、Flink 分布式计算框架，各类 Flink Source API 和 Flink Sink API 操作，以及 Flink 实践的设计（开发）的基本理念、方法与思路。		
实训项目	《实时电商数据分析》项目： 完全由企业工程师负责完成 生成实时数据集，自选角度，设计合理的模型，完成实时电商数据分析，设计流程包括数据预生产、数据分析、实时测试等。		
课程名称	高性能系统架构	学分	2
教学目标	掌握 Hadoop 大数据体系架构及其组件（ZooKeeper、HDFS、HBase、Hive、Spark 等）的技术原理、工作机制，掌握基于 Hadoop 高性能架构的平台的配置、维护和优化方法，能够为复杂大数据工程问题提供优化方案；能够完成 Hadoop 分布式高可用平台的搭建，能够安装、配置、应用性能监控工具，进行 hadoop 平台监控与维护，能够对 HBase、Hive、Spark、HDFS 进行性能优化。		
教学内容	本课程主要介绍 Hadoop 高可用性与高性能架构及其性能优化，包括 Hadoop 高可用性平台搭建、HBase 安装与配置、Hive 安装与配置、Spark 安装与配置、Zookeeper 安装与配置、hadoop 性能优化、HBase 性能优化、Hive 性能优化、Spark 性能优化。		
实训项目	完全由企业工程师负责完成 Hadoop 高性能系统架构方案设计、构建、监控、优化与管理		

九、主要实践环节

1、课程实践环节

序号	课程代码	课程名称	实验学时
1	B0302103	程序设计基础（Python 语言）	32
2	B0302104	数据库原理与技术	32
3	B0302105	数据结构（Python 语言）	32
4	B0302106	计算机网络	24
5	B0302107	操作系统	24
6	B0302108	大数据概论	32
7	B0302209	科技写作	16
8	B0302110	面向对象程序设计（Java 职业认证）	32
9	B0302111	大数据技术架构	24
10	B0302112	分布式计算框架	24
11	B0302113	分布式实时计算	24
12	B0302114	高性能系统架构	16
13	B0302115	数据挖掘	24
14	B0302116	应用统计与试验设计	24
15	B0302117	大数据认证培训	24
16	B0302118	数据可视化	16
合计	400 学时		

2、集中实践环节

序号	课程代码	集中实践环节名称	学分	周数
1	B0302136	数据获取与预处理实践	3	3
2	B0302237	大数据存储与处理项目实践	3	3
3	B0302238	大数据分析可视化实践	3	3
4	B0302239	数据科学综合项目实训	3	3
5	B0700212	军训	1	1
6	B0000201	社会实践	1	
7	B0000202	顶岗实习（含毕业设计或论文）	26	26
合计			40	

十、课程结构比例

大数据工程技术专业课程结构及学时分配表

课程类型	学时	理论教学学时	实践教学学时	占总学时比例
公共基础必修课程	1048	810	238	30.38%
公共基础选修课程	128	128	0	3.71%
专业基础课程	432	240	192	12.52%
专业核心课程	416	208	208	12.06%
专业选修课程	256	128	128	7.42%
集中实践环节	1170		1170	33.91%
合计	3450	1514	1936	100.00%
理论、实践占比		43.88%	56.12%	

集中实践环节学时=（集中实践环节学分-6）*30

大数据工程技术专业课程结构及学分分配表

课程类型	学分	理论学分	实践学分	占总学分比例
公共基础必修课程	61.5	50.5	11	33.70%
公共基础选修课程	8	8	0	4.38%
专业基础课程	27	15	12	14.79%
专业核心课程	26	13	13	14.25%
专业选修课程	16	8	8	8.77%
集中实践环节	40		40	21.92%
第二课堂	4		4	2.19%
合计	182.5	94.5	88	100.00%
理论、实践占比		51.78%	48.22%	

十一、教学计划表

附表1 教学周安排表

附表2 软件工程技术专业教学计划表