

高校大数据专业（专科） 人才培养方案

南京云创大数据科技股份有限公司
Nanjing Innovative Data Technologies, Inc.

2019年03月

计算机类（大数据应用方向） 专科人才培养方案

一、招生对象及学制

（一）招生对象：全日制普通中学高中毕业生，招生方式为普通高考招生。

（二）学习年限：基本学制三年，实行弹性学制，学生在校时间原则上不能少于两年，总在校时间（含休学）不得超过六年。

二、培养目标

培养德、智、体、美全面发展的，能够较快适应生产、建设、管理、服务等一线岗位需要的，面向电信、零售、银行、金融、政府等部门的云平台管理、大数据技术应用与分析的相关工作岗位，具有大数据技术应用与云计算理论知识，掌握云平台运维的基本技能，掌握大数据存储、清洗、管理、建模和分析的基本技能，了解大数据技术应用框架与其生态系统，掌握 Hadoop 技术框架的管理和使用，具有较高综合素质与良好职业素养，能从事云计算运维与大数据技术应用相关基础工作的发展型、复合型、创新型技术技能人才。

三、人才培养规格

（一）基本素质

思想政治素质：掌握马克思主义科学的世界观、人生观和价值观。有坚定跟着共产党走中国特色社会主义道路的信心和决心，有热爱祖国、服务人民的理想信念。具有社会责任感，能积极践行社会主义核心价值观，拥有能够支撑职业和人生发展的思想政治素质。

职业素质：具有良好的职业态度和职业道德修养，具有正确的择业观和创业观。坚持职业操守，爱岗敬业、吃苦耐劳、诚实守信、办事公道、服务群众、奉献社会；具备从事职业活动所必需的基本能力和管理素质；脚踏实地、严谨求实、勇于创新。

人文素养与科学素质：具有融合传统文化精华、当代中西文化潮流的宽阔视野；文理交融的科学思维能力和科学精神；具有健康、高雅、勤勉的生活工作情趣；具有适应社会核心价值体系的审美立场和方法能力；奠定个性鲜明、善于合作的个人成长成才的素质基础。

身心素质：具有一定的体育运动和生理卫生知识，养成良好的锻炼身体、讲究卫生的习惯，掌握一定的运动技能，达到国家规定的体育健康标准；具有坚韧不拔的毅力、积极乐观的态度、良好的人际关系、健全的人格品质。

（二）职业通用能力

1) 具有英语阅读和一般专业资料的翻译能力，通过高职高专应用英语能力 A 或 B 级考试，并获得证书。

2) 通过计算机技术与软件专业技术资格（水平）考试，并获得初级或中级证书。

3) 通过对云计算和大数据技术应用理论知识的学习和实践，能够掌握云计算平台的运维操作与大数据技术应用。

（三）职业岗位能力

1) **大数据处理员岗位能力：**具备大数据应用理论知识，了解大数据技术框架和生态系统，具备 Hadoop 技术框架基础知识，具备程序设计能力，了解非关系型数据库理论知识，具备多数据源整合的基本能力，掌握数据库查询、修改、统计等操作，能对大数据进行预处理、检验和清洗。

2) **大数据分析员岗位能力：**具备大数据应用理论知识，了解大数据技术框架和生态系统，熟悉 Hadoop 技术框架操作，具备程序设计能力，掌握数据库理论知识，掌握数据库增、查、删、改、统计等操作，了解数据仓库与商务智能理论知识，熟悉 ETL 工具的使用，掌握制定报表基本技术，能够进行常规的统计分析、报表分析，能够提出问题，并能制定详细实施计划。

3) **大数据平台运维岗位能力：**掌握 Linux 操作系统理论知识，熟悉

Linux 操作系统操作，掌握 HDFS 理论知识，掌握计算机网络理论知识，具备组建计算机网络、管理网络的能力，了解 Hadoop 及大数据生态系统理论知识，熟悉 Hadoop 技术框架，具备程序设计能力，掌握数据库基础理论知识，具备数据库的基本操作能力。

4) 云平台运维岗位能力：掌握 Linux 操作系统理论知识，熟悉 Linux 操作系统操作，掌握云计算理论知识，具备云计算平台实践的能力，掌握计算机网络理论知识，具备组建计算机网络、管理网络的能力，了解云计算存储与云计算安全理论知识，具备程序设计能力，掌握数据库基础理论知识，具备数据库的基本操作能力。

四、毕业标准

本专业学生在毕业审查时，要求同时达到以下条件：

- (一) 具有良好的思想道德和身体素质。
- (二) 完成培养方案规定的全部教学环节，最低修满 139 学分(其中必修 125 学分)，毕业设计(论文)答辩合格，方可准予毕业并授予学士学位。
- (三) 取得下列人才认证证书之一：
 - 1) Silicon Stone Education (SSE) 大数据人才认证证书。
 - 2) 工业和信息化部颁发的《工业和信息化领域急需紧缺人才培养工程证书》。
 - 3) 软件专业技术水平(资格)考试程序员或软件工程师证书。
 - 4) 其他本专业对应的国家、行业、社会、企业中(初)级且经本学院认定的职业技能证书。

五、课程设置和学分要求

课程类别与学分结构表

课程 模块	必修课		选修课		合计		占总 学分
	学分	学时	最低	最低学	学分	学时	

类别		(周)	学分	时(周)		(周)	比例(%)
公共基础课程	40	704+3周	0	0	40	704+3周	28.78
专业基础课程	33	528	14	224	47	784	33.82
专业核心课程	34	496+4周	0	0	34	496+4周	24.46
毕业设计 与就业指导	18	32+16周	0	0	18	32+16周	12.94
总计	125	1760	14	224	139	1984+23周	100

公共基础课程：包括思想道德修养与法律基础、马克思主义基本原理、毛泽东思想和中国特色社会主义理论体系概论、形势与政策、安全与国防教育、心理健康教育、公共英语、高等数学、线性代数、公共体育等。

专业基础课程：包括数据结构、数据库基础、电路与模拟电子技术、数字电子技术、计算机网络基础、操作系统、计算思维和数据科学、信息论与编码、综合程序设计、Java 语言程序设计、Linux 系统与应用、信息和网络安全等。

专业核心课程：包括云计算导论、大数据导论、大数据实践、大数据系统维护、数据清洗、数据挖掘基础、R 语言、Python 语言、大数据应用实训等。

六、教学进度规划（含主要实践性教学环节、专业实验）

第一学期				
课程名称	课程性质	学分	学时(周)	其中实践学时
军训	必修	2	3周	3周
思想道德修养与法律基础	必修	2	32	
形势与政策	必修	2	32	

公共英语（一）	必修	4	64	
高等数学（一）	必修	4	64	
公共体育（一）	必修	0.5	32	
数据结构	必修	3	48	16
计算机网络基础	必修	2	32	8
Java 语言程序设计	必修	4	64	32
Linux 操作系统	必修	3	48	24
说明：第一学期必修 26.5 学分，最低选修 0 学分，合计 26.5 学分。				

第二学期				
课程名称	课程性质	学分	学时(周)	其中实践学时
马克思主义基本原理	必修	2	32	
安全与国防教育	必修	2	32	
公共英语（二）	必修	4	64	
高等数学（二）	必修	4	64	
公共体育（二）	必修	0.5	32	
操作系统	必修	3	48	16
数据库基础	必修	3	48	24
云计算导论 (专业核心课程)	必修	4	64	16
大数据导论 (专业核心课程)	必修	4	64	16
说明：第二学期必修 26.5 学分，最低选修 0 学分，合计 26.5 学分。				

第三学期				
课程名称	课程性质	学分	学时(周)	其中实践学时
毛泽东思想和中国特色社会主义理	必修	2	32	

论体系概论				
公共英语（三）	必修	4	64	
线性代数	必修	3	48	
公共体育（三）	必修	0.5	32	
电路与模拟电子技术	必修	3	48	16
数字电子技术	必修	3	48	16
计算思维和数据科学	必修	2	32	
信息论与编码	必修	2	32	
Python 语言 （专业核心课程）	必修	4	64	32
大数据实践 （专业核心课程）	必修	4	64	32
说明：第三学期必修 27.5 学分，最低选修 0 学分，合计 27.5 学分。				

第四学期				
课程名称	课程性质	学分	学时(周)	其中实践 学时
心理健康教育	必修	2	32	
公共体育（四）	必修	0.5	32	
综合程序设计	必修	3	48	48
数据挖掘基础 （专业核心课程）	必修	4	64	32
大数据系统运维 （专业核心课程）	必修	3	48	24
数据清洗 （专业核心课程）	必修	3	48	24
计算机程序设计 C	选修	3	48	24
面向对象程序设计(C++)	选修	3	48	24
WEB 程序设计	选修	3	48	24

移动应用开发	选修	3	48	24
说明：第四学期必修 15.5 学分，最低选修 6 学分，合计 21.5 学分。				

第五学期				
课程名称	课程性质	学分	学时(周)	其中实践学时
信息和网络安全	必修	3	48	16
R 语言 (专业核心课程)	选修	4	64	32
大数据应用实训 (专业核心课程)	必修	4	4 周	4 周
项目管理	选修	3	48	24
软件工程	选修	3	48	24
数据建模实训	选修	3	48	48
数据挖掘设计	选修	3	48	48
并行计算与分布式计算	选修	2	32	8
社交网络与舆情分析	选修	2	32	8
说明：第五学期必修 11 学分，最低选修 8 学分，合计 19 学分。				

第六学期				
课程名称	课程性质	学分	学时(周)	其中实践学时
就业指导与职业发展	必修	2	32	
毕业设计	必修	16	16 周	16 周
说明：第六学期必修 18 分，最低选修 0 学分，合计 18 学分。				

七、专业核心课程教材推荐

大数据作为新兴产业，如何实现教学与企业需求相吻合具有重要意义，南京云创大数据科技股份有限公司作为深耕云计算、大数据行业多年的企业，在云计算、大数据方面拥有丰富的实际项目经验和独到的行业见解。同时，经过与各高校多年的深入合作，南京云创大数据科技股份有限公司也更清楚高校育人与企业用人如何有效对接，故对相关专业课程的教材进行推荐，教材的绝大部分内容也是源于企业实际项目，更具有实践意义。

推荐的每本教材皆有配套的 PPT、视频、操作手册、源代码及原始数据，教师教学能更加轻松顺畅，学生也更能体会到实际企业项目的过程，提升教学质量。

（一）专业课程教材推荐表

课程名称	学时数	推荐教材
云计算导论	64	《云计算导论》由刘鹏教授作为丛书总主编率领团队编写，清华大学出版社出版。
大数据导论	64	《大数据导论》由刘鹏教授作为丛书总主编率领团队编写，清华大学出版社出版。
大数据实践	64	《大数据实践》由刘鹏教授作为丛书总主编率领团队编写，清华大学出版社出版。
大数据系统维护	48	《大数据系统运维》由刘鹏教授作为丛书总主编率领团队编写，清华大学出版社出版。
数据清洗	48	《数据清洗》由刘鹏教授作为丛书总主编率领团队编写，清华大学出版社出版。
数据挖掘基础	64	《数据挖掘基础》由刘鹏教授作为丛书总主编率领团队编写，清华大学出版社出版。
R 语言	64	《R 语言》由刘鹏教授作为丛书总主编率领团队编写，清华大学出版社出版。
Python 语言	64	《Python 语言》由刘鹏教授作为丛书总主编率领团队编写，清华大学出版社出版。
大数据应用实训	4 周	《大数据实验手册》由刘鹏教授作为主编率领

		团队编写，电子工业出版社出版。
--	--	-----------------

(二) 推荐教材内容介绍

1. 《大数据导论》



《大数据导论》是了解和学习大数据的基础条件，通过本书了解大数据基本概念，大数据的架构，大数据的采集方式和预处理，常用的 ETL 工具，简单熟悉数据仓库的构建模式，大数据的存储，数据挖掘的方法，以及大数据的可视化技术，从而更好的将大数据技术应用在各行业领域，更深入地开展大数据技术的应用研究。从基础开始，通过理论与实际案例相结合，帮助读者由浅入深进行学习，逐步清理大数据的核心技术和发展趋势。本书可以作为培养应用型人才的课程教材，也适用于初学入门者，对大数据基础理论有需求的广大读者。

2. 《云计算导论》



本书主要内容包括云计算的基本概念、发展现状、主要平台的部署及关键技术、虚拟化与容器技术、云计算的实用化、国内外云计算服务与大规模应用、环境云(envicloud.cn)和万物云(wanwuyun.com)典型行业应用介绍与剖析等内容，

适用于应用型本科、高职高专院校的云计算课程和教学。本书的实验环境部署通过云创大数据实验平台 (<https://bd.cstor.cn>) 上远程开展。

3. 《大数据实践》



本书内容涵盖了目前使用广泛的大数据处理系统 Hadoop 生态圈中的几大核心软件系统：分布式大数据处理系统 Hadoop、Hadoop 数据库 HBase、数据仓库工具 Hive、内存大数据计算框架 Spark 和 Spark SQL，详细介绍了它们的架构、工作原理、部署方法、常用配置、常用操作命令、SQL 引擎等内容。本书对上述几大系统的各种安装部署方式都给出了详细步骤，常用命令也都有具体示例介绍，是一本实操性很强的工具书，能帮助初学者快速掌握和操作这几款常用的大数据处理系统。本书以浅显易懂的语言风格和图文并茂的操作示例引领读者迈入大数据实践之门，可以作为培养应用型人才的课程教材，也可作为相关开发人员的自学教材和参考手册。

4. 《数据清洗》



本书系统地讲解了数据清洗理论和实际应用，共分为 8 章：第 1 章主要介绍数据清洗的概念、任务和流程，数据标准化概念及数据仓库技术等；第 2 章主要

介绍 Windows 和类 UNIX 操作系统下的数据常规格式、数据编码及数据类型转换等；第 3 章介绍 ETL 概念、数据清洗的技术路线、ETL 工具及 ETL 子系统等；第 4 章介绍了 Excel、Kettle、OpenRefine、DataWrangler 和 Hawk 的安装及使用等；第 5 章介绍 Kettle 下文本文件抽取、Web 数据抽取、数据库数据抽取及增量数据抽取等；第 6 章介绍数据清洗步骤、数据检验和数据错误处理，数据质量评估及数据加载；第 7 章介绍网页结构，利用网络爬虫技术进行数据采集，利用 JavaScript 技术进行行为日志数据采集等；第 8 章介绍 RDBMS 的数据清洗方法和数据脱敏处理技术等。

5. 《数据挖掘基础》



本书介绍了数据挖掘的基本概念，包括数据挖掘的常用算法、常用工具、用途和应用场景及应用状况，讲述了常用数据挖掘方法，如分类、聚类、关联规则的概念、思想、典型算法、应用场景等。此外，本书还从实际应用出发，讲解了基于日志的大数据挖掘技术的原理、工具、应用场景和成功案例。通过以上内容的学习，读者将了解数据挖掘的基本概念、思想和算法，并掌握其应用要领。本书可以作为培养应用型人才的课程教材，也可作为相关开发人员的自学教材和参考手册。

6. 《大数据系统运维》



本书是大数据应用人才培养系列教材中的一册，讲解了大数据系统运行维护过程中的各个主要阶段及其任务，包括配置管理、系统管理、故障管理、性能管理、安全管理、高可用性管理、应用变更管理、升级管理及服务资源管理，内容全面且翔实，兼具基础理论知识与运维实践经验，特别是重点介绍了大数据系统的运维特点及运维技能，以保障大数据系统的稳定可靠运行，更好地支撑大数据的商业应用价值。本书具有很强的系统性和实践指导性，可以作为培养应用型人才的课程教材，也同样适合于有意从事 IT 系统运维工作的广大从业者和爱好者作为参考书。

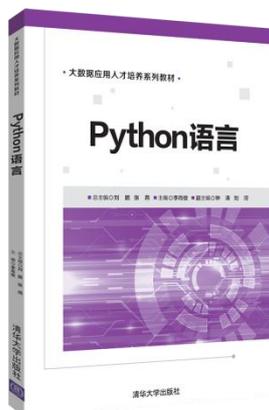
7. 《R 语言》



近年来，R 语言可谓是数据分析的热门语言，相关的资料五花八门，让读者无所适从，本书力求用简洁、精炼、理论与实践相结合的方式让大家快速掌握 R 语言。全书共 14 章，分为基础篇(第 2-10 章)，应用篇(第 11-12 章)和进阶篇(第 13-14 章)。基础篇按照数据分析过程，主要讨论了 R 数据结构、数据导入/导出、数据清洗、数据变换、可视化、高级语言编程和常用建模方法。应用篇通过对 5 个经典案例的分析，使读者能够把学到的 R 基础知识应用到解决实际问题，把数

据变成价值。进阶篇解决如何用 R 处理大数据的一些技术。本书可以作为培养应用型人才课程教材，也可作为数据分析爱好者的参考资料。

8. 《Python 语言》



本书以 WIN10 和 Python3.6.5 搭建 Python 开发基础平台为起点，重点阐述 Python 语言的基础知识和三个典型的项目实战案例。全书以理论引导、案例驱动、上机实战为理念打造 Python 语言学习的新模式。具体内容分为两大部分：第一部分以 Python 编程语言基础知识普及为主，分别介绍了 Python3 概述、基本语法、流程控制、Python 组合数据类型、字符串与正则式、函数、模块、类和对象、异常处理、文件操作；第二部分：以项目实战为核心，以学以致用为导向，以切近生活的案例为依托，分别介绍 Python 爬虫项目实战、Python 数据可视化项目实战、Python 数据分析项目实战。

9. 《大数据实验手册》



《大数据实验手册》内涵 HDFS 实验、YARN 实验、MapReduce 实验、Hive 实验、Spark 实验、ZooKeeper 实验、HBase 实验、Storm 实验、MongoDB 实验、LevelDB 实验、Mahout 实验和综合实战等 42 个大数据实验，每个实验呈现详细的实验目

的、实验内容、实验原理和实验流程指导，涵盖原理验证、综合应用、自主设计及创新的多层次实验内容。本书所有实验均在向公众开放的大数据实验平台（<https://bd.cstor.cn>）和 BDRack 大数据实验一体机上验证通过。

八、培训老师配备与要求

（一）计算机相关专业大学本科及以上学历，5 年或以上工作经验。

（二）3 年以上的大数据行业工作经验，并多次参与大数据项目管理或实施，具备丰富的项目经验。

（三）接受过大数据、软件开发、数据库等方面专业级培训，获得过相关的技能认证证书。

（四）从事过大数据相关系统开发的优先，有过大数据培训课程授课经验的优先。

九、成立专业教学指导委员会

专业教学指导委员会的成员主要来自学校领导、授课老师和行业技术专家。成立专业教学指导委员会的主要工作任务包括：

- 1) 组织和开展本专业教学领域的理论与实践研究。
- 2) 指导本学科专业建设、教材建设、教学改革、实训基地建设、实验室建设等工作。
- 3) 制定专业教学规范。
- 4) 制定教学质量标准。
- 5) 组织师资培训、教学研讨和信息交流等工作。

联系方式:

地址：南京市白下高新技术产业园中国云计算创新基地 A 栋 9 层
电话：400-8855-360 传真：025-83708922
官方网站：<http://www.cstor.cn> 微信公众号：cStor_cn