

大数据科学与技术专业 人才培养方案(本科)

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

2019年03月

大数据科学与技术专业 本科人才培养方案

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) 标准学制：标准学制为 4 年，学习年限 3-6 年，授予工学学士学位。

二、培养目标

“大数据科学与技术”瞄准社会各领域对大数据高级应用型人才的需求，致力于培养具有扎实的信息科学、大数据科学知识，熟练掌握大数据采集、处理、分析与应用的技术与工具，能够承担企业、事业、政府、社会组织等部门的信息管理、信息咨询服务、信息研究等工作，具有大数据系统设计、实现与应用能力的专业技术人才。

本专业培养大数据科学与工程领域的复合型高级工程技术人才。毕业生品格健全，具有科学的人文精神、创新创业精神和良好的职业道德精神，具备自主学习能力、批判思维能力和国际交流能力。毕业生具有信息科学、管理科学和数据科学基础知识与基本技能，掌握大数据科学与技术所需要的计算机、网络、数据编码、数据处理等相关学科的基本理论和基本知识，熟练掌握大数据采集、存储、处理与分析、传输与应用等技术，具备大数据工程项目的系统集成能力、应用软件设计和开发能力，具有一定的大数据科学研究能力及数据科学家岗位的基本能力与素质。

毕业后能从事各行业大数据分析、处理、服务、开发和利用工作，大数据系统集成与管理维护等各方面工作，亦可从事大数据研究、咨询、教育培训工作。

三、人才培养规格

本专业培养掌握大数据科学与技术相关的基本理论和基本知识，系统地掌握数据科学与工程专业知识，具备大数据应用系统设计与开发的能力，以及一定的科研工作能力，达到知识、能力与素质的协调发展。

毕业生在知识、能力和素质等方面应达到如下具体要求：

1) 知识要求：具有良好的自然科学基础，扎实的信息科学基础；具有较好的人文社会科学、管理科学知识；熟练掌握大数据科学与技术核心专业知识和应用技术，主要包括信息论与编码、计算思维和数据科学、云计算、大数据、深度学习、数据挖掘、大数据数据库技术、大数据可视化、Hadoop 系统开发、大数据应用实训、虚拟化与容器等。

2) 能力要求：具备从事大数据应用系统设计与实现的能力，特别在数据分析、数据管理、数据存储等方面，受到较系统的工程训练，能发现、分析和解决实际工程技术问题。具备良好的工程项目交流、表达、组织、管理、协调与沟通的能力。了解信息学科、计算机学科、数据科学的发展动态，并掌握相关文献检索方法，具有基本的专业资料分析与综合的能力，良好的文档与科学论文撰写能力。具有较强的创新意识，一定的创新创业能力，

3) 素质要求：毕业生身心健康；有良好的道德修养，尊重生命、遵纪守法、诚信友善、乐于奉献；有高尚的民族精神，积极弘扬传统文化，热爱祖国，崇尚集体主义精神；有坚定的理想信念，拥护中国特色社会主义，贯彻科学发展观、和谐社会理论和“四个全面”思想。

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 184 学分(其中必修 144 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 类别	必修课		选修课		合计		占总 学分 比例(%)
	学分	学时 (周)	学分	学时 (周)	学分	学时 (周)	
公共基 础课程	52	928+3 周	0	0	52	928+3 周	28.26
专业基 础课程	43	544	40	640	83	1184	45.11
专业核 心课程	32	480+6 周	0	0	32	480+6 周	17.39
毕业设 计与就 业指导	17	18周	0	0	17	18周	9.24
总计	144	1952+27 周	40	640	184	2592+27 周	100%

公共基础课程：包括马克思主义基本原理、思想道德修养与法律基础、毛泽东思想和中国特色社会主义理论体系概论、中国近现代史纲要、形势与政策、大学英语、高等数学、大学物理、线性代数、离散数据、概率论与数理统计、体育、大学生心理健康教育等。

专业基础课程：包括大数据科学与技术导论、计算机程序设计基础C、面向

对象程序设计(C++)、操作系统原理、信息论与编码、计算思维和数据科学、数据结构、数据库原理、计算机网络、电路与模拟电子技术、数字电子技术、计算机原理与汇编等。

专业核心课程：包括云计算、大数据、Hadoop 系统开发、大数据库、大数据应用实训、深度学习、数据挖掘、虚拟化与容器、大数据可视化、Python 程序设计等。

六、教学进度安排（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
军事理论课	必修	1	16	
体育(一)	必修	1	32	
思想道德修养与法律基础	必修	3	48	
形势与政策	必修	1	16	
大学英语(一)	必修	4	64	
高等数学(一)	必修	4	64	
大数据科学与技术导论(新生课)	必修	1	16	
计算机程序设计基础C	必修	3	48	24
大学物理	必修	4	64	16
线性代数	必修	2	32	
说明：第一学期必修 26 学分，最低选修 0 学分，合计 26 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	3	48	
大学生心理健康教育	必修	2	32	
体育(二)	必修	1	32	
大学英语(二)	必修	4	64	
高等数学(二)	必修	4	64	
概率论与数理统计	必修	3	48	
电路与模拟电子技术	必修	4	64	16
面向对象程序设计(C++)	必修	3	48	24
程序设计综合训练(上)	必修	2	32	32
说明：第二学期必修 26 学分，最低选修 0 学分，合计 26 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
大学英语(三)	必修	4	64	
体育(三)	必修	1	32	
数据结构	必修	3	48	16
数字电子技术	必修	3	48	16
电子技术实验	必修	1	32	32
信息论与编码	必修	3	48	
计算思维和数据科学	必修	3	48	
云计算 (专业核心课程)	必修	3	48	16
Java 语言程序设计	选修	2	32	32
理工学术英语	选修	2	32	
说明：第三学期必修 21 学分，最低选修 2 学分，合计 23 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
体育(四)	必修	1	32	
计算机网络	必修	3	48	16
计算机原理与汇编	必修	3	48	16
数据库原理	必修	3	48	16
程序设计综合训练(下)	必修	3	3周	3周
大数据 (专业核心课程)	必修	3	48	16
数据挖掘 (专业核心课程)	必修	3	48	24
算法分析与设计	选修	3	48	16
操作系统原理	选修	3	48	
运筹学	选修	2	32	
电子技术课程设计	选修	2	32	32
英语公众演讲	选修	2	32	16
说明：第四学期必修 19 学分，最低选修 5 学分，合计 24 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
体育课外测试(一)	必修	1		
离散数学	必修	3	48	
大数据库 (专业核心课程)	必修	3	48	24
大数据可视化 (专业核心课程)	必修	3	48	24
Hadoop 系统开发 (专业核心课程)	必修	3	48	24
大数据课程设计	选修	2	32	32

云计算课程设计	选修	2	32	32
数据结构课程设计	选修	2	32	32
大数据与领域建模	选修	2	32	16
并行计算与分布式计算	选修	2	32	8
网络流量监测与分析	选修	2	32	8
多源信息融合	选修	2	32	8
推荐算法及应用	选修	2	32	16
说明：第五学期必修 15 分，最低选修 10 学分，合计 25 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
体育课外测试(二)	必修	1		
中国近现代史纲要	必修	2	32	
深度学习 (专业核心课程)	必修	3	48	24
Python 程序设计 (专业核心课程)	必修	4	64	32
文档聚类与查重技术	选修	2	32	16
R 语言与网络数据获取技术	选修	3	48	24
Linux 系统与应用	选修	3	48	24
软件工程	选修	3	48	24
多媒体原理与系统设计	选修	2	32	8
计算机视觉	选修	2	32	8
Web 技术与网络资源管理	选修	3	48	24
移动应用开发	选修	3	48	24
社交网络与舆情分析	选修	2	32	16
说明：第六学期 必修 10 学分，最低选修 13 学分，合计 23 学分。				

第七学期				
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论体系概论	必修	3	48	
信息与网络安全	必修	3	48	16
虚拟化与容器 (专业核心课程)	必修	3	48	24
大数据应用实训 (专业核心课程)	必修	4	6周	6周
生物信息学	选修	2	32	8
数据可视化分析	选修	2	32	16
商务智能及应用	选修	2	32	8
计算机仿真与建模	选修	2	32	8
人机交互	选修	2	32	8
数据分析与融合	选修	2	32	8
数据可视化课程设计	选修	2	32	32
大数据综合应用课程设计	选修	2	32	32
说明：第七学期必修 13 学分，最低选修 10 学分，合计 23 学分。				

第八学期				
课程名称	课程性质	学分	学时(周)	其中实践学时
毕业设计	必修	16	16周	16周
毕业教育	必修	1	2周	
说明：第八学期必修 17 分，最低选修 0 学分，合计 17 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）专业核心课程教材推荐表

课程名称	学时数	推荐教材
云计算	48	《云计算》由刘鹏教授作为主编率领团队编写，电子工业出版社出版。
大数据	48	《大数据》由刘鹏教授作为主编率领团队编写，电子工业出版社出版。
Hadoop 系统开发	48	《实战 Hadoop2.0 从云计算到大数据》由刘鹏教授作为主编率领团队编写，电子工业出版社出版。
大数据库	48	《大数据库》由刘鹏教授作为主编率领团队编写，电子工业出版社出版。
大数据应用实训	6 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。
深度学习	48	《深度学习》由刘鹏教授作为主编率领团队编写，电子工业出版社出版。
数据挖掘	48	《数据挖掘》由刘鹏教授作为丛书总主编率领团队编写，电子工业出版社出版。
虚拟化与容器	48	《虚拟化与容器》由刘鹏教授作为丛书总主编率领团队编写，电子工业出版社出版。

大数据可视化	48	《大数据可视化》由刘鹏教授作为丛书总主编率领团队编写，电子工业出版社出版。
Python 程序设计	64	《Python 程序设计》由刘鹏教授作为丛书总主编率领团队编写，电子工业出版社出版。

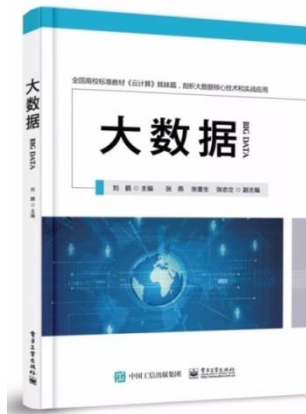
(二) 推荐教材内容介绍

1. 《云计算》



《云计算》第一版于2010年3月出版，第二版于2011年5月出版，时隔4年，第三版于2015年7月出版。《云计算》已经成为了全国高校首选教材，素有云计算领域“红宝书”之称，内容主要包括大数据与云计算，Google 云计算原理与应用，Amazon 云计算 AWS，微软云计算 Windows Azure，Hadoop 2.0 主流开源云架构，Hadoop 2.0 大家族，虚拟化技术，OpenStack 开源虚拟化平台，云计算数据中心，云计算核心算法，中国云计算技术。

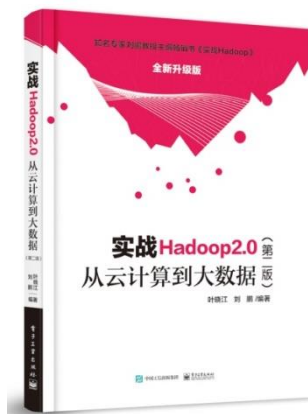
2. 《大数据》



《大数据》是国内绝大多数高校采用的知名教材《云计算》(1~3版)的姊妹

妹篇，是中国大数据专家委员会刘鹏教授联合国内多位专家历时两年的心血之作。大数据领域一直缺乏一本权威教材，希望本书能够填补空白。本书系统地介绍了大数据的理论和实战应用，包括大数据采集与预处理、数据挖掘算法与工具、深度学习以及大数据可视化等，并深度剖析了大数据在互联网、商业和典型行业的应用。

3. 《实战 Hadoop2.0 从云计算到大数据》



《实战 Hadoop》第一版于 2011 年 9 月出版，是国内第一本 Hadoop 编程书籍，经过 5 年的精心准备，第二版于 2016 年 6 月出版。本书内容主要包括：大数据组件概述，大数据集群，集群管理工具 Ambari，分布式文件系统 HDFS，分布式资源管理器 YARN，分布式并行处理 MapReduce，分布式锁服务 ZooKeeper，分布式数据库 HBase，内存型计算框架 Spark，数据流实时处理系统 Storm，数据仓库工具 Hive，其他大数据组件，手工部署 Hadoop2.0。

4. 《大数据库》



《大数据库》系统地介绍了大数据库的理论和实战应用，包括大数据库概述、分布式数据库 HBase、数据仓库工具 Hive、大数据查询系统 Impala、内

存数据库 Spark、Spark SQL、键值数据库、流式数据库和数据应用托管平台 Docker 等，并紧跟大数据的发展前沿，既有理论深度，又有使用价值。通过本书可掌握大数据库的概念和原理，学习主要的大数据库技术和核心算法。

5. 《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。

6. 《深度学习》



深度学习开启了人工智能发展的新一轮高潮。《深度学习》为没有任何机器学习基础的人提供了一个深度学习的入门教材，使用浅显易懂的语言和深入浅出的方式循循善诱，给初学者一个从原理到实践的深度学习大餐。本书内容主要包括：深度学习来源及应用，深度学习的数学基础，人脑神经系统与深度学习，深

深度学习基本过程，深度学习主流模型，深度学习开源框架，深度学习在图像中的应用，深度学习在语音中的应用，深度学习在文本中的应用，深度学习前沿发展。

7. 《数据挖掘》



《数据挖掘》的定位是大数据挖掘技术与应用。通过本书可掌握经典数据挖掘理论、方法、工具与应用，大数据环境下常用数据挖掘算法的优化，大数据新常态下催生的数据分析方法（如推荐系统、链接分析与网页排序、互联网信息抽取、日志挖掘与查询分析）、工具与应用等。

8. 《虚拟化与容器》



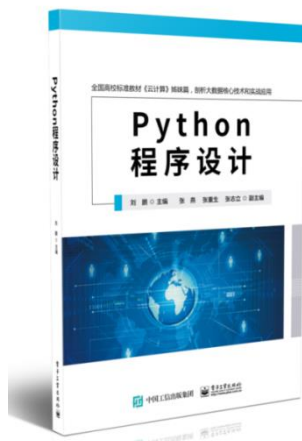
《虚拟化与容器》系统介绍了从虚拟化技术到容器技术发展的历史必然性及其发展趋势。其中针对虚拟化技术详细阐述了包括计算、存储及网络虚拟化技术的原理、架构及相关案例；针对容器技术详细阐述了容器技术内核原理，以及 Docker、LXC 和 Kubernetes 容器技术核心原理及相关案例。

9. 《大数据可视化》



有图有真相；有树木，有森林；沙中淘金，是大数据可视化的价值。《大数据可视化》教材在分析视觉特性基础上，尝试建立大数据理论与实践应用之间桥梁，读者能够体验数据之美，科学之美。作为大数据应用的分支，可视化有其自己的逻辑，本书尝试在大数据应用之海，擦亮可视化的贝壳，为大数据应用锦上添花。

10. 《Python 程序设计》



本书是一本全面的从入门到实践的 Python 编程教程，从带领读者快速掌握基本的编程知识开始，循序渐进、层层升入的引导读者利用新学到的知识开发功能丰富的项目。全书首先介绍了 Python 的基本概念和安装；随后介绍了包括数据类型、文件操作、程序调试、类和对象、连接数据源等基本知识；并结合网络爬虫、数据挖掘、自然语言处理、数据可视化、Web 和移动应用等工具，以案例为依托进行项目实战；最后介绍了国内各种云服务平台，以及如何运用 Python 实现访问。本书的全部实验均可在大数据实验平台(<https://bd.cstor.cn/login>)上远程开展，也可在高校部署的 BDRack 大数据实验一体机上本地开展。

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) 5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