



# 《大数据技术原理与应用（第3版）》

<http://dblab.xmu.edu.cn/post/bigdata3>

温馨提示：编辑幻灯片母版，可以修改每页PPT的厦大校徽和底部文字

## 课程介绍

（PPT版本号：2020年12月版本）

林子雨 博士/副教授

厦门大学计算机科学系

E-mail: [ziyulin@xmu.edu.cn](mailto:ziyulin@xmu.edu.cn) ▶▶

主页： <http://dblab.xmu.edu.cn/post/linziyu>





# 本章配套教学视频

《大数据技术原理与应用（第3版）》  
在线视频观看地址

<http://www.icourse163.org/course/XMU-1002335004>

大数据技术原理与应用

**BIGDATA TECHNOLOGY AND APPLICATION**

打开大数据之门，遨游大数据世界





# 提纲

- 1.课程特色
- 2.教材介绍
- 3.内容提要
- 4.教学大纲
- 5.配套资源
- 6.班级介绍

附录A：主讲教师林子雨简介

附录B：《大数据技术原理与应用》教材

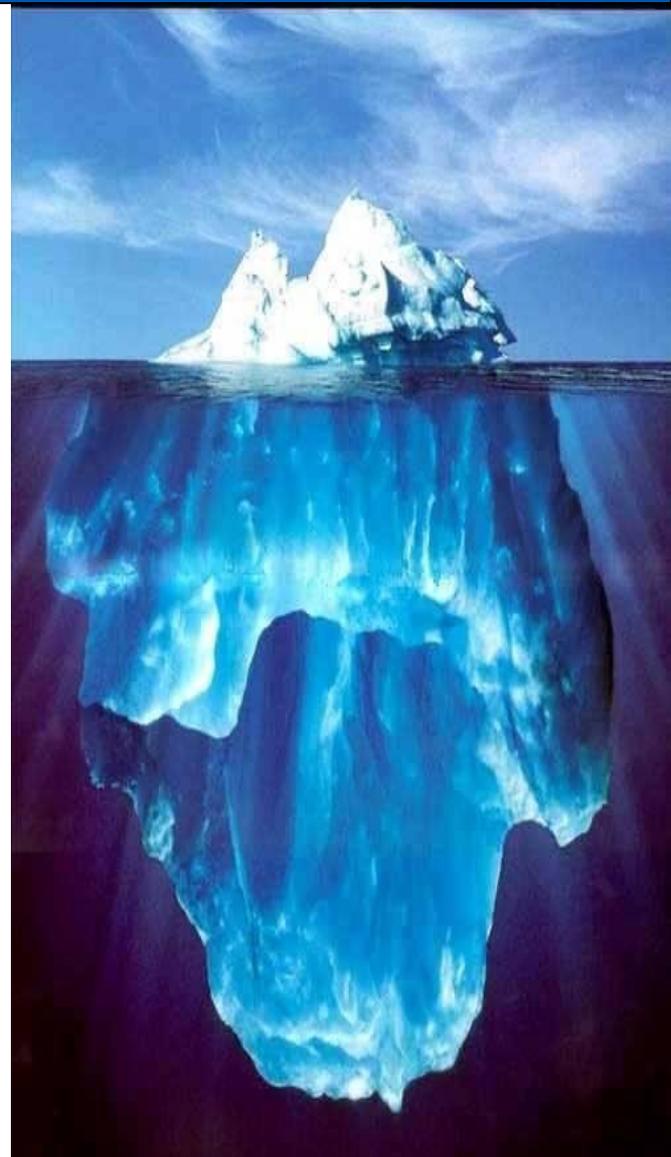
附录C：高校大数据课程公共服务平台

本PPT是如下教材的配套讲义：

《大数据技术原理与应用  
——概念、存储、处理、分析与应用》  
(2020年9月第3版)

厦门大学 林子雨 编著，人民邮电出版社

欢迎访问《大数据技术原理与应用（第3版）》教材官方网站：  
<http://dbl原因lab.xmu.edu.cn/post/bigdata3>





# 主讲教师



2017年度厦门大学奖教金获得者  
2020年度厦门大学奖教金获得者

主讲教师：厦门大学 林子雨 博士/副教授  
中国高校首个“数字教师”提出者和建设者

2009年7月从事教师职业以来  
累计**免费**网络发布超过**1000万**字高价值教学和科研资料  
网络浏览量超过**1000万**次



# 课程特色



搭建起通向“大数据知识空间”的桥梁和纽带  
构建知识体系、阐明基本原理  
引导初级实践、了解相关应用  
为学生在大数据领域“深耕细作”奠定基础、指明方向



# 教材介绍

## 《大数据技术原理与应用——大数据概念、存储、处理、分析与应用》

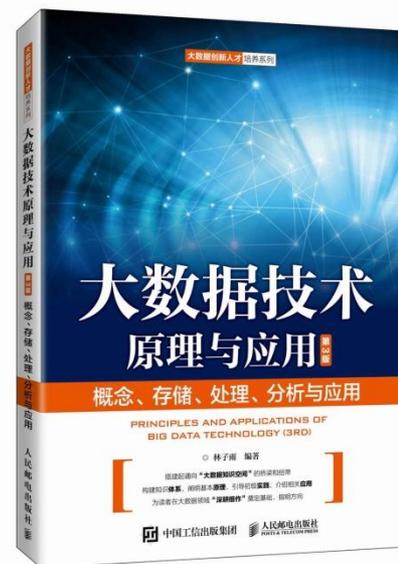
厦门大学 林子雨编著 人民邮电出版社 2021年1月第3版

ISBN:978-7-115-54405-6 定价：59.8元

- 国内高校第一本系统介绍大数据知识专业教材
- 京东、当当等各大网店畅销书籍
- 致力于打造成为大数据入门教材精品
- 工信部“全国云计算与大数据应用技术人才考试、认证项目”目前唯一指定大数据教材
- 国内多所高校采用本教材开课
- 配套目前国内高校最完备的课程公共服务平台



扫一扫  
访问教材官网



2021年1月第3版



# 教材配套实验手册

本书是与《大数据技术原理与应用（第3版）》教材配套的唯一指定实验指导书

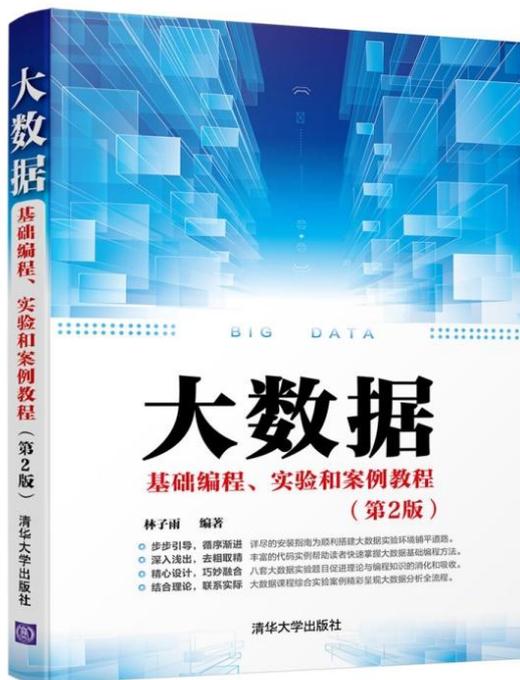
林子雨编著《大数据基础编程、实验和案例教程（第2版）》

清华大学出版社 ISBN:978-7-302-55977-1 定价：69元，2020年10月第2版

教材官网：<http://dbllab.xmu.edu.cn/post/bigdatapRACTICE2/>



扫一扫访问  
教材官网



- 步步引导，循序渐进，详尽的安装指南为顺利搭建大数据实验环境铺平道路
- 深入浅出，去粗取精，丰富的代码实例帮助快速掌握大数据基础编程方法
- 精心设计，巧妙融合，八套大数据实验题目促进理论与编程知识的消化和吸收
- 结合理论，联系实际，大数据课程综合实验案例精彩呈现大数据分析全流程



# 内容提要

- 本课程系统介绍了大数据相关知识
- 系统地论述了大数据的基本概念、大数据处理架构Hadoop、分布式文件系统HDFS、分布式数据库HBase、NoSQL数据库、云数据库、分布式并行编程模型MapReduce、数据仓库Hive、基于内存的分布式计算框架Spark、流计算、流处理框架Flink、图计算以及大数据在互联网、生物医学和物流等各个领域的应用
- 在Hadoop、HDFS、HBase、MapReduce、Hive、Spark和Flink等重要章节，安排了入门级的实践操作，让学生更好地学习和掌握大数据关键技术



# 教学大纲

章节	主要内容	学时安排
第1章 大数据概述	介绍大数据的基本概念和应用领域，并阐述大数据、云计算和物联网的相互关系	2
第2章 大数据处理架构Hadoop	介绍大数据处理架构Hadoop	4
第3章 分布式文件系统HDFS	分布式文件系统HDFS的基本原理和使用方法	4
第4章 分布式数据库HBase	分布式数据库HBase的基本原理和使用方法	4
第5章 NoSQL数据库	NoSQL数据库的概念和基本原理	2
第6章 云数据库	云数据库的概念和基本原理	2
第7章 MapReduce	分布式并行编程模型MapReduce原理和使用方法	4
第8章 Hadoop架构再探讨	Hadoop技术的发展演变过程、YARN、HDFS Federation	自学
第9章 数据仓库Hive	数据仓库Hive的概念、原理和安装使用方法	2
第10章 Spark	Spark原理与基础编程	4
第11章 流计算	流计算概念和原理、Storm	2
第12章 Flink	流处理框架Flink的应用场景、体系架构和编程模型	2
第13章 图计算	图计算概念和原理、Pregel	自学
第14章 数据可视化	数据可视化概念、代表性技术和软件	自学
第15、16和17章 大数据在不同领域的应用	大数据在互联网领域的典型应用：推荐系统，以及在其他领域的典型应用	自学
合计		32



# 教材官网

- 教材官网: <http://dbllab.xmu.edu.cn/post/bigdata3/>
- 提供以下教材配套内容:
- 讲义PPT
- 授课视频
- 教材配套大数据软件安装和编程实践指南
- 实验指南和答案
- 习题（选择题）和答案



# 高校大数据课程公共服务平台

所有资料全部免费共享  
支持电脑和手机浏览



## 高校大数据课程 公共服务平台



<http://dblab.xmu.edu.cn/post/bigdata-teaching-platform/>



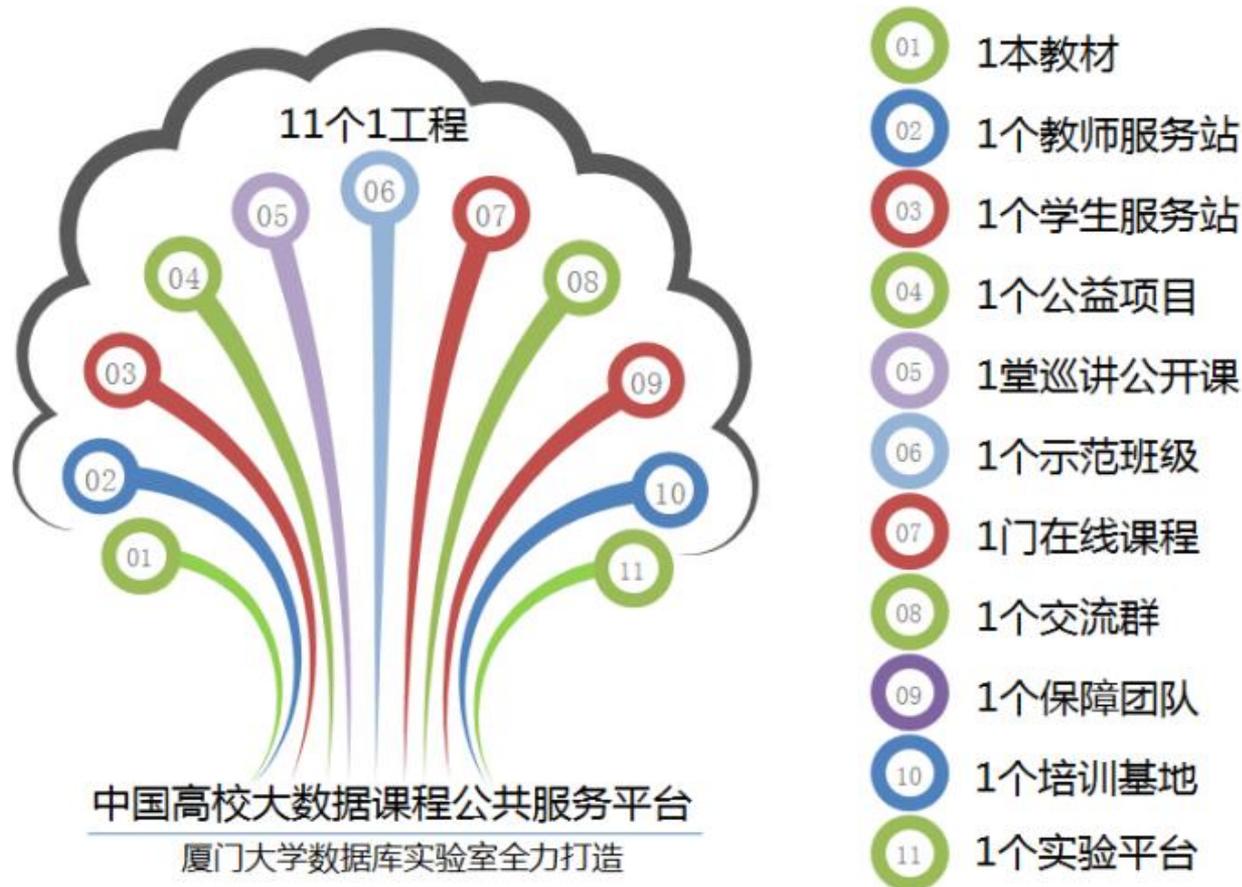
扫一扫访问平台主页



扫一扫观看3分钟FLASH动画宣传片



# 课程配套教学资源



- 国内高校第一个大数据课程公共服务平台
- 为教师教学和学生学习大数据课程，提供全方位、一站式服务
- 平台所有资源全部通过网络**免费共享（无需注册）**、采用WORD和PPT可编辑格式



# 课程配套教学资源

制作精美的平台宣传图册





<http://dblab.xmu.edu.cn/post/xuanchuanpian/>

中国高校大数据课程公共服务平台

# 3分钟FLASH 动画宣传片



# 附录A：主讲教师林子雨简介



## 主讲教师：林子雨

单位：厦门大学计算机科学系

E-mail: [ziyulin@xmu.edu.cn](mailto:ziyulin@xmu.edu.cn)

个人网页: <http://dblab.xmu.edu.cn/post/linziyu>

数据库实验室网站: <http://dblab.xmu.edu.cn>



扫一扫访问个人主页

林子雨，男，1978年出生，博士（毕业于北京大学），全国高校知名大数据教师，现为厦门大学计算机科学系副教授，曾任厦门大学信息科学与技术学院院长助理、晋江市发展和改革局副局长。中国计算机学会数据库专业委员会委员，中国计算机学会信息系统专业委员会委员。国内高校首个“数字教师”提出者和建设者，厦门大学数据库实验室负责人，厦门大学云计算与大数据研究中心主要建设者和骨干成员，2013年度、2017年度和2020年度厦门大学教学类奖教金获得者，荣获2019年福建省精品在线开放课程、2018年厦门大学高等教育成果特等奖、2018年福建省高等教育教学成果二等奖、2018年国家精品在线开放课程。主要研究方向为数据库、数据仓库、数据挖掘、大数据、云计算和物联网，并以第一作者身份在《软件学报》《计算机学报》和《计算机研究与发展》等国家重点期刊以及国际学术会议上发表多篇学术论文。作为项目负责人主持的科研项目包括1项国家自然科学基金青年基金项目(No.61303004)、1项福建省自然科学基金青年基金项目(No.2013J05099)和1项中央高校基本科研业务费项目(No.2011121049)，主持的教改课题包括1项2016年福建省教改课题和1项2016年教育部产学协作育人项目，同时，作为课题负责人完成了国家发改委城市信息化重大课题、国家物联网重大应用示范工程区域试点泉州市工作方案、2015泉州市互联网经济调研等课题。中国高校首个“数字教师”提出者和建设者，2009年至今，“数字教师”大平台累计向网络免费发布超过1000万字高价值的研究和教学资料，累计网络访问量超过1000万次。打造了中国高校大数据教学知名品牌，编著出版了中国高校第一本系统介绍大数据知识的专业教材《大数据技术原理与应用》，并成为京东、当当网等网店畅销书籍；建设了国内高校首个大数据课程公共服务平台，为教师教学和学生学习大数据课程提供全方位、一站式服务，年访问量超过200万次，累计访问量超过1000万次。





# 附录C：林子雨大数据系列教材



林子雨大数据系列教材

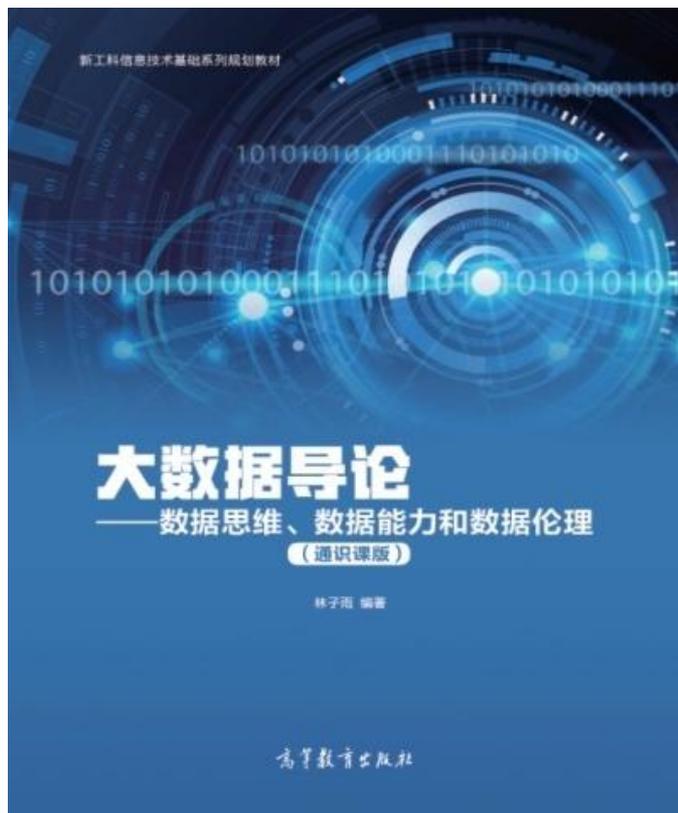
用于导论课、专业课、实训课、公共课

了解全部教材信息：<http://dbllab.xmu.edu.cn/post/bigdatabook/>



# 附录D：《大数据导论（通识课版）》教材

## 开设全校公共选修课的优质教材



本课程旨在实现以下几个培养目标：

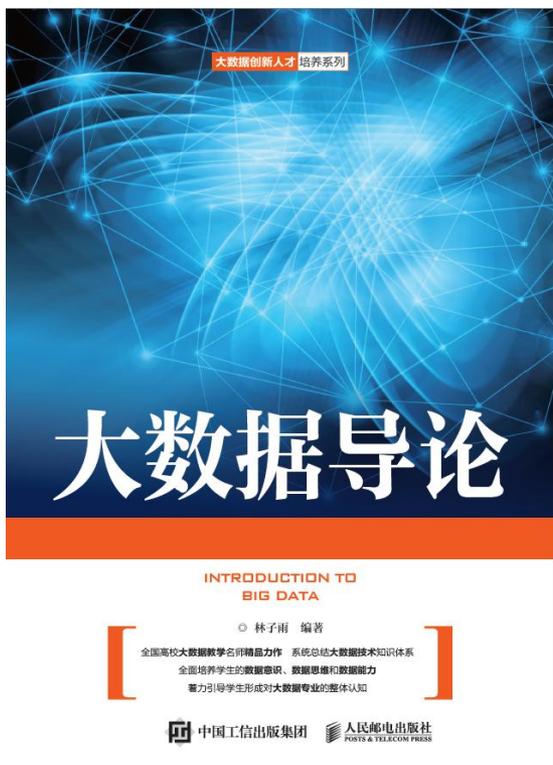
- 引导学生步入大数据时代，积极投身大数据的变革浪潮之中
- 了解大数据概念，培养大数据思维，养成数据安全意识
- 认识大数据伦理，努力使自己的行为符合大数据伦理规范要求
- 熟悉大数据应用，探寻大数据与自己专业的应用结合点
- 激发学生基于大数据的创新创业热情

高等教育出版社 ISBN:978-7-04-053577-8 定价：32元 版次：2020年2月第1版  
教材官网：<http://dbl原因.xmu.edu.cn/post/bigdataintroduction/>



# 附录E：《大数据导论》教材

- 林子雨 编著 《大数据导论》
  - 人民邮电出版社，2020年9月第1版
  - ISBN:978-7-115-54446-9 定价：49.80元
- 教材官网：<http://dbl原因.xmu.edu.cn/post/bigdata-introduction/>



开设大数据专业导论课的优质教材



扫一扫访问教材官网



# 附录F：《大数据技术原理与应用（第3版）》教材

《大数据技术原理与应用——概念、存储、处理、分析与应用（第3版）》，由厦门大学计算机科学系林子雨博士编著，是国内高校第一本系统介绍大数据知识的专业教材。人民邮电出版社 ISBN:978-7-115-54405-6 定价：59.80元

全书共有17章，系统地论述了大数据的基本概念、大数据处理架构Hadoop、分布式文件系统HDFS、分布式数据库HBase、NoSQL数据库、云数据库、分布式并行编程模型MapReduce、Spark、流计算、Flink、图计算、数据可视化以及大数据在互联网、生物医学和物流等各个领域的应用。在Hadoop、HDFS、HBase、MapReduce、Spark和Flink等重要章节，安排了入门级的实践操作，让读者更好地学习和掌握大数据关键技术。

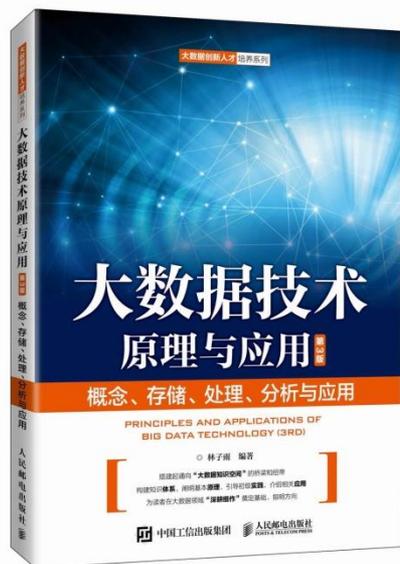
本书可以作为高等院校计算机专业、信息管理等相关专业的大数据课程教材，也可供相关技术人员参考、学习、培训之用。

欢迎访问《大数据技术原理与应用——概念、存储、处理、分析与应用》教材官方网站：

<http://dmlab.xmu.edu.cn/post/bigdata3>



扫一扫访问教材官网





# 附录G：《大数据基础编程、实验和案例教程（第2版）》

本书是与《大数据技术原理与应用（第3版）》教材配套的唯一指定实验指导书

大数据教材



1+1黄金组合  
厦门大学林子雨编著

配套实验指导书



- 步步引导，循序渐进，详尽的安装指南为顺利搭建大数据实验环境铺平道路
- 深入浅出，去粗取精，丰富的代码实例帮助快速掌握大数据基础编程方法
- 精心设计，巧妙融合，八套大数据实验题目促进理论与编程知识的消化和吸收
- 结合理论，联系实际，大数据课程综合实验案例精彩呈现大数据分析全流程

林子雨编著《大数据基础编程、实验和案例教程（第2版）》

清华大学出版社 ISBN:978-7-302-55977-1 定价：69元 2020年10月第2版



# 附录H：《Spark编程基础（Scala版）》

## 《Spark编程基础（Scala版）》

厦门大学 林子雨，赖永炫，陶继平 编著

披荆斩棘，在大数据丛林中开辟学习捷径  
填沟削坎，为快速学习Spark技术铺平道路  
深入浅出，有效降低Spark技术学习门槛  
资源全面，构建全方位一站式在线服务体系

人民邮电出版社出版发行，ISBN:978-7-115-48816-9  
教材官网：<http://dblalab.xmu.edu.cn/post/spark/>

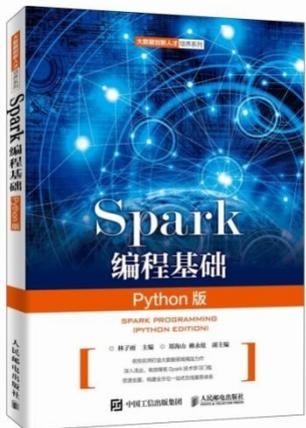


本书以Scala作为开发Spark应用程序的编程语言，系统介绍了Spark编程的基础知识。全书共8章，内容包括大数据技术概述、Scala语言基础、Spark的设计与运行原理、Spark环境搭建和使用方法、RDD编程、Spark SQL、Spark Streaming、Spark MLlib等。本书每个章节都安排了入门级的编程实践操作，以便读者更好地学习和掌握Spark编程方法。本书官网免费提供了全套的在线教学资源，包括讲义PPT、习题、源代码、软件、数据集、授课视频、上机实验指南等。



# 附录I: 《Spark编程基础 (Python版)》

## 《Spark编程基础 (Python版)》



厦门大学 林子雨, 郑海山, 赖永炫 编著

披荆斩棘, 在大数据丛林中开辟学习捷径  
填沟削坎, 为快速学习Spark技术铺平道路  
深入浅出, 有效降低Spark技术学习门槛  
资源全面, 构建全方位一站式在线服务体系

人民邮电出版社出版发行, ISBN:978-7-115-52439-3

教材官网: <http://dblab.xmu.edu.cn/post/spark-python/>



本书以Python作为开发Spark应用程序的编程语言, 系统介绍了Spark编程的基础知识。全书共8章, 内容包括大数据技术概述、Spark的设计与运行原理、Spark环境搭建和使用方法、RDD编程、Spark SQL、Spark Streaming、Structured Streaming、Spark MLlib等。本书每个章节都安排了入门级的编程实践操作, 以便读者更好地学习和掌握Spark编程方法。本书官网免费提供了全套的在线教学资源, 包括讲义PPT、习题、源代码、软件、数据集、上机实验指南等。



# 附录J：高校大数据课程公共服务平台



## 高校大数据课程

公 共 服 务 平 台

<http://dbllab.xmu.edu.cn/post/bigdata-teaching-platform/>



扫一扫访问平台主页



扫一扫观看3分钟FLASH动画宣传片



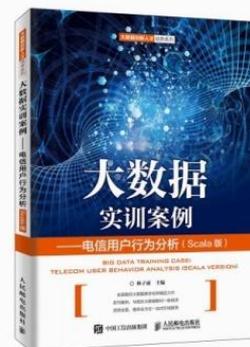
# 附录K：高校大数据实训课程系列案例教材

为了更好地满足高校开设大数据实训课程的教材需求，厦门大学数据库实验室林子雨老师团队联合企业共同开发了《高校大数据实训课程系列案例》，目前已经完成开发的系列案例包括：

- 《电影推荐系统》（已经于2019年5月出版）
- 《电信用户行为分析》（已经于2019年5月出版）
- 《实时日志流处理分析》
- 《微博用户情感分析》
- 《互联网广告预测分析》
- 《网站日志处理分析》



系列案例教材将于2019年陆续出版发行，教材相关信息，敬请关注网页后续更新！  
<http://dblab.xmu.edu.cn/post/shixunkecheng/>



扫一扫访问大数据实训课程系列案例教材主页

The background of the slide features several faint, light-blue silhouettes of people. At the top, there are two groups of people standing and holding hands. On the right side, a person is shown in profile, looking towards the center. On the left side, two people are shown in profile, one appearing to be speaking or gesturing towards the other. The overall scene suggests a group of people in a meeting or presentation setting.

**Thank You!**

**Department of Computer Science, Xiamen University, 2020**