

课程简介

- 操作性强，理论性强、难度适中
- 相关行业就业的基础
- 课程性质：选修课
- 学时：**32学时**

课程目的和基本要求

- 课程性质
 - 技术基础
- 基础知识要求
 - JAVA，数据结构，数据库原理/ORACLE，
 - LINUX，计算机系统/体系结构，计算机网络
- 主要特点
 - 既有理论，又有实践
 - 面向应用设计
 - 涉及数据的组织

大数据人才

- **大数据系统研发工程师**: 大规模非结构化数据业务模型构建、大数据存储、数据库构设、优化数据库构架
- **大数据分析师**: 搭建大数据应用平台以及开发分析应用程序, 熟悉工具或算法、编程、优化及部署不同**MapReduce**, **ETL**开发
- **大数据应用开发工程师**: 数据挖掘工作, 运用算法来解决和分析问题, 具备**Hadoop**框架经验
- **大数据可视化工程师**: 利用图形化工具揭示数据中的复杂信息, 新型数据可视化工具如**Spotifre**, **Qlikview**和**Tableau**
- **数据安全研发工程师**: 负责企业内部大型服务器、存储、数据安全管理工作, 并对网络、信息安全项目进行规划、设计和实施
- **数据科学专家**

课程目的和基本要求

- 三种基本的能力
 1. 大数据的思维能力
 2. 常用工具软件和算法
 - 3. 大数据常规分析算法与模型**

课程要求

- 多方查阅资料
- 无故不得旷课，有事要请假
- 按时独立地完成作业
- 实验课之前要认真准备，认真、独立做实验

希望

勤于思考、乐于实践



大数据技术及应用

课程介绍

马新娜

计算机科学系

E-mail: maxinnamxn@163.com

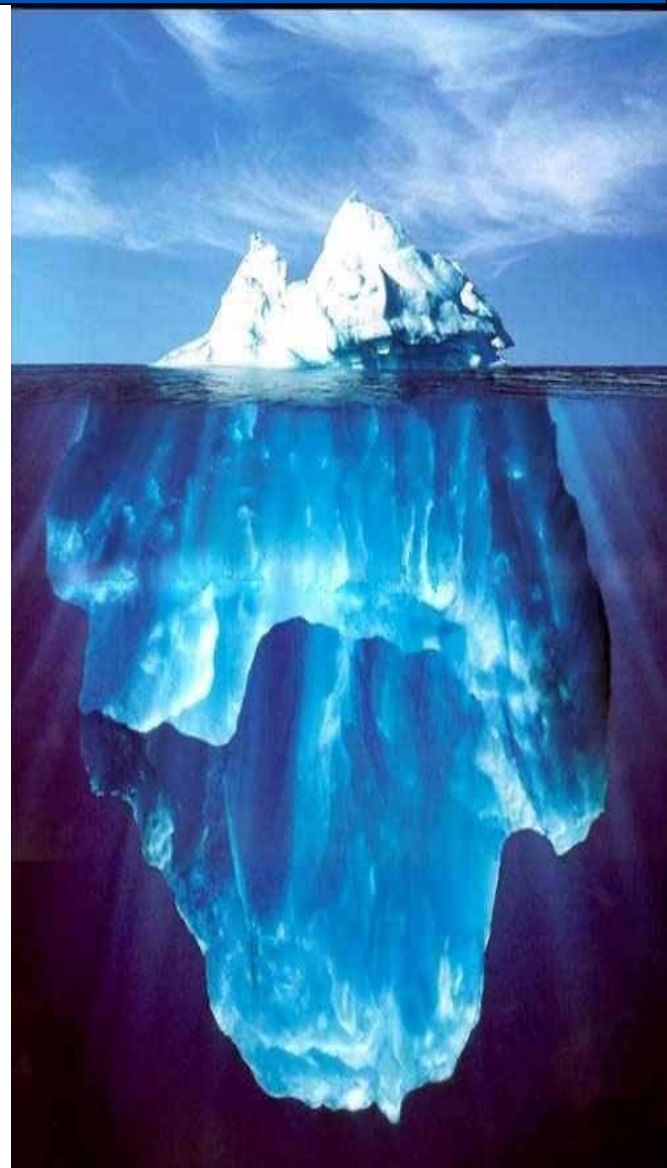
提纲

1.课程特色

2.教材介绍

2.内容提要

3.篇章安排



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，2018年2月第2版第7次印刷
21世纪高等教育计算机规划教材 ISBN:978-7-115-39287-9

内容简介：

- （1）概念篇：介绍当前紧密关联的最新IT领域技术云计算、大数据和物联网。
- （2）大数据存储篇：介绍分布式数据存储的概念、原理和技术，包括HDFS、HBase、NoSQL数据库、云数据库。
- （3）大数据处理与分析篇：介绍MapReduce分布式编程框架、图计算、流计算。
- （4）大数据应用篇：介绍基于大数据技术的推荐系统。

篇章安排

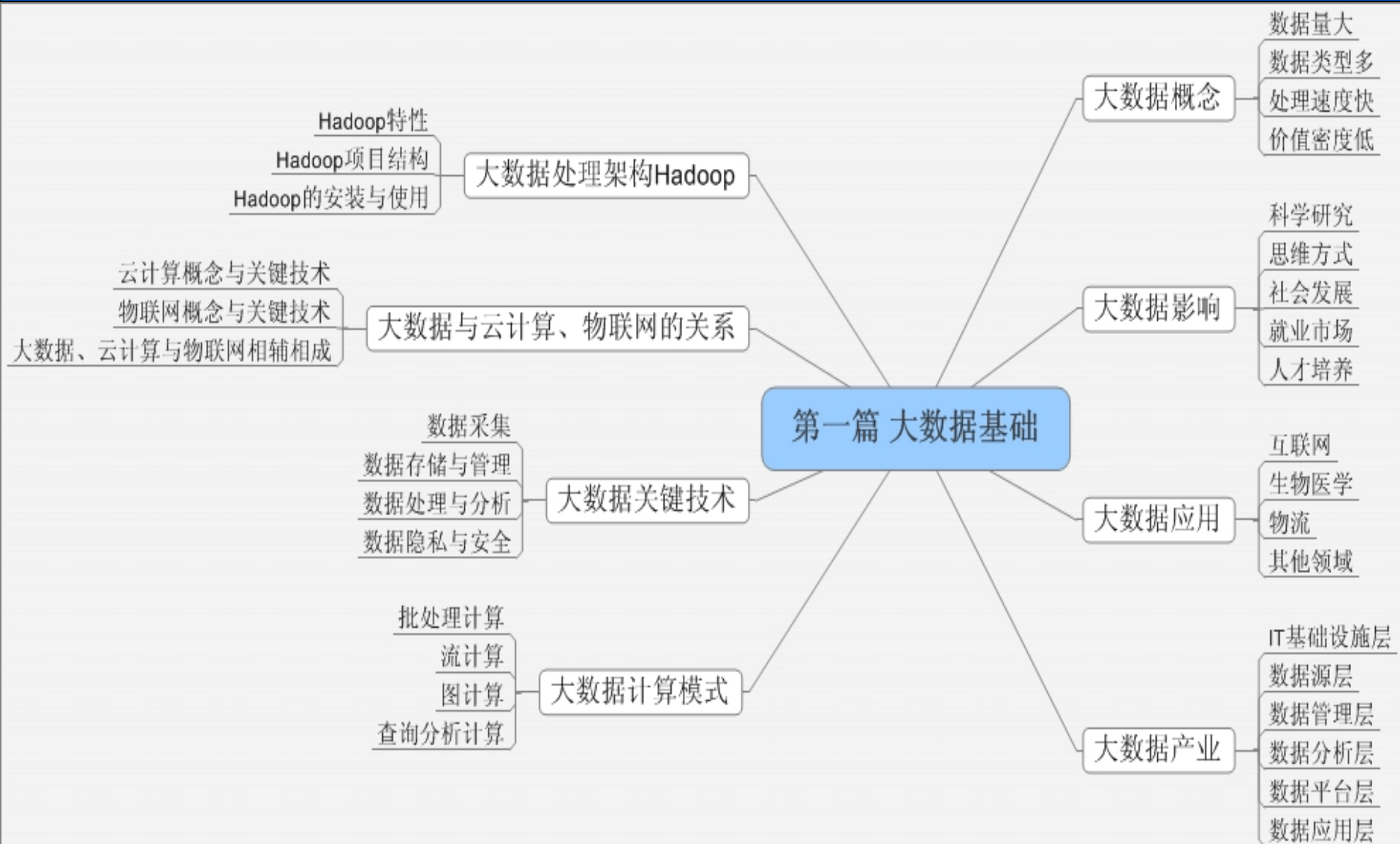
第一篇：大数据基础篇

第二篇：大数据存储篇

第三篇：大数据处理与分析篇

第四篇：大数据应用篇

第一篇：大数据基础篇



第二篇：大数据存储篇

第二篇 大数据存储

分布式文件系统HDFS

- HDFS体系结构
- HDFS存储原理
- HDFS编程实践

分布式数据库HBase

- HBase数据模型
- HBase实现原理
- HBase运行机制
- HBase编程实践

NoSQL数据库

- NoSQL数据库概念

- NoSQL数据库四大类型

- 键值数据库
- 列式数据库
- 文档数据库
- 图形数据库

- NoSQL数据库三大基石

- CAP
- BASE
- 最终一致性

- 从NoSQL到NewSQL数据库

云数据库

- 云数据库概念
- 云数据库产品

- 云数据库系统架构UMP

- UMP系统概述
- UMP系统架构
- UMP系统功能

- 云数据库实践

- 阿里云RDS

第三篇：大数据处理与分析篇

第三篇 大数据处理与分析

静态数据：批处理

分布式并行编程框架MapReduce

MapReduce工作流程

MapReduce具体应用

MapReduce编程实践

流数据：实时计算

流计算框架Storm

流计算处理流程

Storm设计思想

Storm实例

图结构数据

图计算框架Pregel

Pregel图计算模型

Pregel体系结构

Pregel应用实例

数据可视化

入门级工具

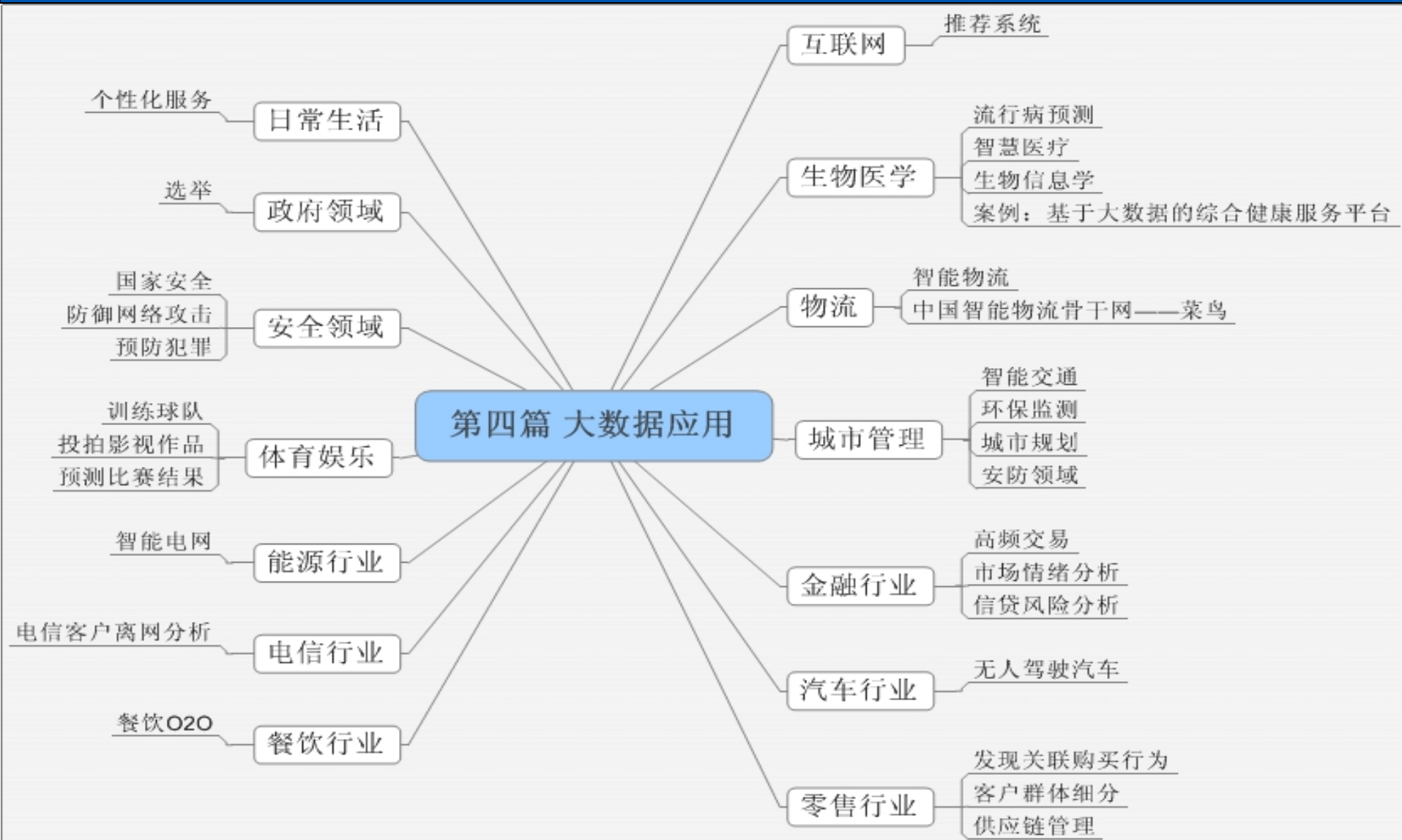
信息图表工具

地图工具

时间线工具

高级分析工具

第四篇：大数据应用篇



各章内容

第一章 大数据概述

第二章 大数据处理架构Hadoop

第三章 分布式文件系统HDFS

第四章 分布式数据库HBase

第五章 NoSQL数据库

第六章 云数据库

第七章 MapReduce

第八章 流计算

第九章 图计算

第十章 可视化

第十一章 应用

大数据分析应用