第一章 绪论

1. **教学目标**
2. 了解本课程的研究对象、内容；
3. 了解本课程的学习目的、考核方式等；
4. 理解大数据的发展现状。
5. **教学内容**
6. 课程的研究对象
7. 课程的学习目的、考核方式
8. 大数据前沿发展和选题介绍等
9. **教学重点与难点**
10. 本章的教学重点是大数据的相关概述，学生能够了解大数据的研究对象、主要研究内容，以及相关领域应用；
11. 如何准确理解大数据的内涵和外延，特别是大数据的一般处理过程，包括数据获取、数据管理、数据分析以及数据可视化和数据治理等；
12. 如何运用大数据相关知识、原理和技术，解决学术界和产业界所面临的不同复杂科学和工程问题。
13. **教学方法与手段**

本章是课程中前沿概述章节，对本章内容的掌握需要教师首先从宏观层面深入浅出的阐述大数据的相关知识、原理和技术，通过学术界的前沿科学问题和产业界的紧迫工程难题，以案例的方式，让学生对大数据有一个整体的理解和掌握。

本章以教师讲授（多媒体+板书）为主，师生交互、学生讨论为辅的教学方法与手段进行教学。

1. **教学过程**

* 课时1

（1）教师通过具体的案例，请同学们参与讨论与自己息息相关的大数据实际问题和应用场景。并请同学们思考大数据的一般处理过程。

（2）引出大数据的主要研究对象、主要研究内容，并针对性的给出大数据的行业应用。

（3）请同学们思考并举例讨论：大数据到底有多大？大数据对自己的生活带来了什么变化？

（4）教师讲授大数据的作用和意义，分别从经济、社会、科学等方面，让学生理解数据新经济时代的核心竞争力。

* 课时2

（1）教师介绍大数据的技术挑战与科学意义，并结合一个具体例子，阐述大数据的处理过程。

（2）教师通过具体的案例，请同学们参与讨论与自己息息相关的大数据实际问题和应用场景。并请同学们思考大数据的一般处理过程。

（3）教师讲授数据科学的概念，以及学术和产业视角。

1. **课后作业**
2. 当前大数据技术上面临的主要挑战是什么？
3. 请阐述4V和3I分别是什么含义。

第二章 大数据基本概念

1. **教学目标**

1. 了解大数据的多边定义、价值期望；

2. 了解大数据的生命周期、技术图谱；

3. 理解大数据的应用场景、机遇与挑战。

1. **教学内容**
2. 大数据的多边定义、价值期望
3. 大数据的生命周期、技术图谱
4. 大数据的应用场景、机遇与挑战
5. **教学重点与难点**
6. 本章的教学重点是大数据的基本概述，学生能够了解什么是大数据，以及大数据的技术支撑和主要来源；
7. 理解大数据的特征及意义，以及大数据的主要表现形态；
8. 了解大数据的主要应用场景。
9. **教学方法与手段**

本章是该课程中概念章节，对本章内容的掌握需要教师首先从大数据的发展历史、数据规模的增长以及所引起的学术界和产业界问题为切入点，阐述大数据的基本概念、特征、表现形态和应用场景等，通过学术界和产业界的前言案例介绍，让学生对大数据的相关概念有深入理解和掌握。

本章以教师讲授（多媒体+板书）为主，师生交互、学生讨论为辅的教学方法与手段进行教学。

1. **教学过程**

* 课时1

（1）教师通过数据规模的发展，结合学生学习、工作和生活中所涉及的大数据场景，请同学们参与讨论与自己息息相关的大数据概念。

（2）引出大数据的主要概念。

（3）引出大数据的来源和意义。

* 课时2

（1）教师介绍大数据的主要表现形态。

（2）教师通过具体的案例，讲解大数据的实际应用场景。并请同学们思考在解释实际大数据问题时需要注意的问题和挑战。

（3）教师总结大数据的主要概念，为后续章节的具体知识点讲授奠定认知上的基础。

1. **课后作业**

结合自己的生活经历，举例说明一个大数据实际的应用场景，思考其中所涉及到的技术问题。

第三章 大数据获取与感知

1. **教学目标**
2. 了解数据的来源与类型；
3. 了解数据源的一般分布；
4. 理解数据源分布的多源、异构特性；
5. 掌握不同数据源的获取方法和策略；
6. 掌握数据抽取、转换、装载的过程；
7. 掌握典型数据采集的开源工具和技术。
8. **教学内容**
9. 数据渠道
10. 数据获取方法
11. 数据采集平台
12. **教学重点与难点**
13. 数据的来源与类型；
14. 不同数据源的获取方法和策略；
15. 数据抽取、转换、装载的过程；
16. 典型数据采集的开源工具和技术。
17. **教学方法与手段**

本章是该课程中难度较大的一章，对本章内容的掌握不能仅仅依靠记忆，更需要学生掌握数据获取方法和策略，以及典型数据采集的开源工具和技术等。因此本章的教学须辅以适量练习和讲解。本章最后将布置程序练习，让学生利用开源软件进行数据获取及处理实验。

本章以教师讲授（多媒体+板书）为主，师生交互、学生讨论、实践指导为辅的教学方法与手段进行教学。

1. **教学过程**

* 课时1

（1）请同学回答上一章的一个问题。

（2）引出“数据感知”的问题，指出人们在日常生活中从数据中获取知识的实例。

（3）教师讲授大数据发展情况，说明大数据处理的必要性。

（4）请同学来列举一下日常生活中接触的大数据有哪些，讨论这些大数据来源有什么相同点。

（5）教师对大数据的来源及分布进行总结，并对三种主要来源方式进行讲述。

（6）教师讲授大数据四大特点并进行举例说明。

（7）请同学来对比传统数据与大数据的不同点。教师逐步尝试通过对比大数据与传统数据的不同，请学生讨论对与大数据的获取及处理过程会存在哪些难点。

（8）对大数据目前的难点进行总结，并引出数据源的获取方法和策略的讲解。

* 课时2

（1）教师对行业/企业内数据的获取方法及获取必要性进行讲授。

（2）教师讲授ETL整体流程。

（3）教师讲解数据的抽取、转换与加工、加载等的具体概念及操作。

（4）老师对学生进行现场提问，掌握同学的学习情况，并根据回答确定是否进行再次讲解。问题：在ETL转换与加工过程中主要需要进行的操作。

（5）教师讲授ETL主流开源工具，重点讲述Kettle工具。

* 课时3

（1）教师讲授互联网发展以及互联网数据获取的必要性。

（2）教师讲授互联网数据的特点。

（3）教师讲授互联网数据获取的方式-网络爬虫的概念、流程。

（4）教师讲授网络爬虫URL的抓取策略，包括IP地址搜索策略、广度优先、深度优先、最佳优先、反向链接数、Partial PageRank策略、OPIC策略共7种。

（5）提问学生广度优先和深度优先两种策略的URL抓取顺序，结合学生们回答进行讲解。

（6）教师讲解网络爬虫主流开源工具。

* 课时4

（1）教师介绍物联网概念及其发展历史。

（2）教师讲授物联网大数据来源及其特点。

（3）教师介绍物联网大数据获取的主要方式：RFID、传感器等并进行举例说明。

（4）布置程序练习，学生自己动手利用开源网络爬虫工具进行页面爬取测试。

（5）教师总结全章。

1. **课后作业**

方案编程题；如果给你一个目标网站，你要如何爬取网站的有用内容？阐述主要思路和实现方案。

第四章 大数据存储与管理

1. **教学目标**
2. 掌握分布式文件系统及主流技术HDFS；
3. 掌握分布式数据库中典型技术HBase；
4. 理解非关系型数据库兴起的原因，及其与传统关系型数据库的差异。
5. **教学内容**
6. 分布式文件系统
7. 分布式数据库
8. 非关系型数据库
9. **教学重点与难点**
10. 本章的教学重点是介绍数据存储与管理技术的概念与发展过程，选择经典的关系数据库技术以及大数据时代的分布式文件系统技术、NoSQL与Sql on Hadoop技术新型大数据存储与查询技术进行介绍；
11. 如何运用大数据存储与管理相关知识、原理和技术，解决学术界和产业界所面临的不同复杂科学和工程问题。
12. **教学方法与手段**

本章是该课程中大数据核心技术的关键章节之一，对本章内容的掌握需要教师理论结合实践，以课程讲授、案例剖析以及师生讨论的方式，让学生掌握大数据存储与管理的相关知识、原理和技术。

本章以教师讲授（多媒体+板书）为主，师生交互、学生讨论、实践指导为辅的教学方法与手段进行教学。

1. **教学过程**

* 课时1

（1）教师通过数据管理的发展历史，引出数据管理的内涵。

（2）引出关系模型，让学生理解关系数据库正是建立在关系数据模型之上的，用以存储结构化数据并支持数据的插入、查询、更新、删除等操作。

（3）请同学们思考并举例讨论：关系型数据库以及关系模型的优缺点，特别是在大数据场景下所面临的挑战。

（4）引入分布式文件系统的概念，并介绍相应的几种分类方式。

* 课时2

（1）教师介绍Hadoop分布式文件系统HDFS的功能和特点，并介绍相应解决的实际问题场景。

（2）教师深入介绍Hadoop的架构和操作。

（3）教师讲解Hadoop的副本管理概念和具体技术。

* 课时3

（1）教师介绍Ceph的概念和应用场景。

（2）通过对比，介绍Ceph的优势和核心设计思想。

（3）教师深入介绍Ceph的技术架构。

* 课时4

（1）回顾Hadoop和Ceph相关知识点。

（2）教师介绍GlusterFS的功能和特点，并介绍相应解决的实际问题场景。

（3）教师深入介绍GlusterFS的技术架构。

* 课时5

（1）请同学们思考并讨论：关系型数据库和分布式数据管理之间的对比，以及各自所使用的场景。

（2）教师引入非关系型数据库的机遇，并结合案例阐述非关系型数据库的实际应用场景。

* 课时6

（1）教师介绍NoSQL数据库的功能和特点，并介绍相应解决的实际问题场景。

（2）教师深入介绍NoSQL的技术架构。

（3）教师举例介绍当前主流的NoSQL数据库。

1. **课后作业**

分析设计题：阐述HDFS和HBase在Hadoop生态中的功能和联系，以及未来技术展望。基于华为智能基座的合作，在华为云服务器搭建一个Hadoop集群。

第五章 大数据分析与挖掘

1. **教学目标**
2. 了解常用特征表示、特征提取和特征选择的方法；
3. 理解常用的数据建模方法，掌握关联规则挖掘基本概念和常用算法；
4. 理解常用的数据挖掘算法，能运用于分析和解决实际问题；
5. 了解大数据可视化方法与技术，能应用常见可视化工具进行可视化分析。
6. **教学内容**
7. 数据理解与特征工程
8. 常用数据挖掘算法
9. 数据可视化
10. **教学重点与难点**
11. 让学生需要了解常见的数据类型，掌握数据规范、度量方法、特征工程，其中度量方法和特征工程也是教学的难点。
12. 大数据建模方法，主要内容包括监督学习和非监督学习，理解机器学习的过程和方法。
13. 机器学习中的非监督学习的聚类分析技术，以及聚类算法K均值算法(K-means)，包括算法执行过程、优化函数、收敛条件等。
14. 大数据可视化技术和相关工具，以及如何进行实践。
15. **教学方法与手段**

本章是该课程中重点和难点，对本章内容的掌握不能仅仅依靠记忆，更需要学生深刻理解大数据分析与挖掘技术的概念、掌握大数据建模技术等。因此本章的教学必须辅以大量练习和讲解，并能提供学生一种易于理解和使用的算法设计的基本方式方法。本章最后将布置程序练习，学生自己动手实现聚类分析算法和分类算法。

本章以教师讲授（多媒体+板书）为主，师生交互、学生讨论、实践指导为辅的教学方法与手段进行教学。

1. **教学过程**

* 课时1

（1）请同学回答关于现有的大数据类型。引出数据理解的问题，指出大数据中的“数据”不仅仅指数值数据，还包括文本、图像、网络、时空数据等。

（2）教师讲授数据类型，讲授数据特征、属性等概念，并给出相应的实例。讲授数据集的类型，以及特点、特性等。

（3）教师讲授数据规范的概念和方法。

（4）请同学思考如何度量两个数据的问题，可以采用什么方法。

（5）教师讲授数据特征工程的基本概念。

（6）教师讲授数据特征表示方法和技术。

（7）教师讲授数据特征构建的方法和技术。

（8）教师讲授不同数据的特征提取技术。

* 课时2

（1）教师讲授数据建模技术：机器学习的概念、常用的库、分类方法，包括非监督学习和监督学习。

（2）教师讲授非监督学习中的重要一类——关联规则挖掘。

（3）教师介绍关联规则的基本概念，包括定义、支持度、置信度、频繁项集。同时介绍关联规则挖掘的应用场景。

（4）教师通过商场购买商品的问题——如何分析用户购买商品之间的关联？如何将不同商品进行联合促销？背后的原因？进一步将前面讲授的理论知识进行强化，并引出关联规则挖掘算法。

（5）教师讲授关联规则挖掘过程和Apriori算法。

（6）教师详细讲授Apriori算法，通过实例讲授Apriori算法的算法思想、执行过程、改进思路。

* 课时3

（1）教师通过“物以类聚”的例子引出非监督学习的另一种分析技术——聚类分析。介绍聚类技术的基本概念。

（2）教师讲授聚类算法的算法思想，并回顾聚类算法中数据之间距离或是相似度度量方法，引出K均值算法（K-means）。

（3）教师讲授K均值算法的算法执行过程，是同学们对算法的执行过程有初步的认知。

（4）教师讲授K均值算法伪代码，并展示算法执行的过程。同时从优化学习的角度讲解K均值算法，与学生探讨K均值算法中，K的取值对聚类算法的影响。

（5）教师讲授聚类算法的评价方法。

* 课时4

（1）回顾上次课讲授的非监督学习的内容，引出本节内容——监督学习，讲解监督学习的建模方式。

（2）通过住房面积与房价问题，教师讲授线性回归建模。讲授回归问题的数学表达和模型。

（3）通过线性回归模型讲授机器学习中的重要概念——损失函数和求解目标。

（4）教师重点讲解梯度下降算法，梯度下降算法是机器学习中重要的求解方式。讲授李同梯度下降算法求解线性回归问题，以及线性回归的评估方法。

（5）通过线性回归模型嵌套logistic函数，进而讲授logistic回归对分类问题的建模。

（6）教师讲授分类问题的问题定义，以及logistic回归的求解方法。

（7）教师讲授分类问题的评估方法。

* 课时5

（1）回顾上次课讲授的监督学习的线性回归和logistic回归，引出本节内容——分类，讲解分类问题的建模方式。

（2）通过与朋友用餐的例子，与学生一起探讨，引入KNN算法的基本思想。

（3）教师讲授KNN算法原理、执行步骤。

（4）教师讲授KNN算法中的K值选择与距离度量函数。

（5）通过分类问题，继续引入另一大分类算法朴素贝叶斯，教师讲授朴素贝叶斯的基本概念。

（6）通过统计学习，教师讲授朴素贝叶斯，理解朴素贝叶斯的数学原理、目标函数，求解方法。

（7）教师讲授朴素贝叶斯的计算实例，详细求解步骤。

* 课时6

（1）教师与学生们讨论人类思考和反应的基础——神经元，并启发学生人类是怎么对外界食物进行学习的，如何反应的，机制是怎样的？通过讨论，引入人工神经元的概念，并讲授神经元的数学建模。

（2）教师讲授人工神经网络，以及BP神经网络，梯度反向回传算法求解BP神经网络。

（3）教师讲授卷积神经网络的相关结构以及若干改进。详细讲授卷积操作与池化层。

（4）教师讲授循环神经网络的相关结构以及改进的长段时记忆神经网络(LSTM)，并详细讲授LSTM中的相关结构，包括忘记门、输入门、输出门、更新过程。

（5）教师讲授生成式对抗神经网络（GANs）包括GANs的构成、训练方法、改进的DGANs。

* 课时7

（1）本节是本章的最后一部分内容，大数据可视化的方法与技术，首先通过若干可视化的实例引入大数据可视化，为学生展示大数据可视化的效果。

（2）教师讲授大数据可视化的基本概念、可视化的重要作用、可视化的主要流程等。

（3）教师讲授大数据可视化的方法和技术，讲授不同的数据类型的可视化方法，如文本数据的标签云、网络数据、时空数据、多维数据可视化等。

（4）教师讲授当前的可视化的软件工具，如D3、Echart、Tableau、Gephi等，并简要介绍对应的数据类型和可开发的效果。

（5）教师总结本节和本章内容。

1. **课后作业**

算法编程题：布置两次程序实践的专题训练任务，一次为非监督学习关联规则挖掘和聚类分析技术，一次为监督学习的分类技术，并要求通过数据可视化技术进行展示。

第六章 大数据架构与处理

1. **教学目标**
2. 了解早期流行的集中式分布架构；
3. 掌握当前主流的分布式计算构架：MapReduce和Spark；
4. 了解大数据计算处理加速技术：GPU、TPU和FPGA。
5. **教学内容**
6. 集中式计算架构
7. 分布式计算架构
8. 处理加速技术
9. **教学重点与难点**
10. 本章的教学重点与难点是大数据架构与处理，主要涉及传统的集中式计算架构、分布式计算架构以及大数据处理的硬件加速技术；其中，分布式计算架构是本章的重点和难点；
11. 现在流行的分布式计算架构：Mapreduce，包括其体系结构、工作流程和实例。此外，Spark的框架，生态和运行机理也是本章的另一个重点和难点；

3. 大数据的硬件加速技术：GPU、TPU、FPGA。

1. **教学方法与手段**

本章的内容对于学生理解大数据处理技术的整体软件架构，对本章内容的掌握不能仅仅依靠了解和记忆，更需要学生深刻理解大数据架构设计思想、工作流程、生态系统等。因此本章的教学必须辅以一定量的理论讲解和实例讲解，并能提供学生一种易于理解的大数据架构和处理可视化展示。另外，本章最后将简要介绍大数据硬件加速技术，GPU、TPU和FPGA。

本章以教师讲授（多媒体+板书）为主，师生交互、学生讨论、实践指导为辅的教学方法与手段进行教学。

1. **教学过程**

* 课时1

（1）教请同学讨论计算机硬件发展的情况。

（2）引出大型主机和超级计算机，教师讲授相对于个人PC大型主机的发展、特点和优势。

（3）教师讲授目前大型主机所面临的问题。

（4）教师讲授超级计算机，介绍中国在超级计算机领域的贡献和所取得成就，以及其中的典型代表天河二号。

（5）教师引导学生导论大数据时代，超级计算机所面临的挑战，科学计算与大数据计算的区别与联系。

* 课时2

（1）教师讲授Mapreduce概述。比较传统并行计算框架与Mapreduce在集群架构/容错性、硬件/价格/扩展性、编程/学习难度、适用场景四个维度的不同。

（2）教师讲授Mapreduce的体系结构。其中要重点讲授Mapreduce的理念，Map函数和reduce函数。Mapreduce体系结构中的四部分：Client、JobTracker、TaskTracker、Task，对其中的每一部分进行仔细讲解。

（3）教师Mapreduce工作流程。主要内容有：MapReduce各个执行阶段和Shuffle过程详解。

（4）教师讲授分布式文件系统，包括计算机集群结构和分布式文件系统结构。

（5）教师介绍HDFS。HDFS需要实现的目标：兼容廉价的硬件设备、流数据读写、大数据集、简单的文件模型、强大的跨平台兼容性。以及HDFS的主要组件及其功能。

（6）教师讲授与HDFS相关的一些概念。

* 课时3

（1）教师回顾上节课讲授的Mapreduce和HDFS的内容，随机抽查学生对于大数据的Mapreduce处理机制的掌握情况。

（2）教师讲授HDFS体系结构，包括HDFS体系结构概述、HDFS命名空间管理、通信协议、客户端、HDFS体系结构的局限性。

（3）教师讲授HDFS存储原理冗余数据保存、数据存取策略、数据错误与恢复。对HDFS整体内容进行总结。

（4）教师对Spark进行简要介绍，内容有Spark发展和主要特点。

（5）教师讲授Spark生态系统，包括和HDFS、Mapreduce的联系，Spark的优势等。

（6）教师Spark运行架构，包括架构设计、Spark运行的基本流程，其中RDD是Spark中的重点和难点，教师讲授RDD概念、依赖关系、Stage的划分、RDD运行过程等。

（7）教师讲授流式数据计算架构，Apache Storm，Apache Samza。

* 课时4

（1）教师引导学生讨论实际过程中大数据的数据量是多大，前面提到的机器学习算法、深度学习算法，大数据架构与处理平台面临的数据挑战，与当前的硬件平台是否能够匹配。

（2）教师讲授目前的大数据处理的加速技术概述，和大数据处理中的常见数学操作。

（3）教师讲授图形处理器（GPU）、张量处理器（TPU）、可编程逻辑阵列。

（4）教师总结全章。

1. **课后作业**

分析应用题：调研总结当前国内外大数据处理加速相关技术的最新进展，特别是国内国产自主的相关技术（例如华为的相关产品）。

第七章 大数据安全与隐私保护

1. **教学目标**
2. 了解大数据安全的意义和面临的问题；
3. 了解大数据隐私保护的意义，理解常用的隐私保护技术；
4. 了解大数据在网络安全中的应用。
5. **教学内容**
6. 大数据安全
7. 大数据隐私保护
8. 大数据与网络安全
9. **教学重点与难点**
10. 大数据安全与隐私保护的意义；
11. 大数据隐私保护技术；
12. 大数据在网络安全中的应用。
13. **教学方法与手段**

本章是课程中难度适中的一章，对本章内容的掌握不能仅仅依靠记忆，更需要学生充分了解大数据安全和隐私保护的重要性大以及数据在网络安全中的应用。本章的教学通过课堂提问实时了解学生对于知识点的掌握情况。

本章以教师讲授（多媒体+板书）为主，师生交互、学生讨论、案例分析研讨为辅的教学方法与手段进行教学。

1. **教学过程**

* 课时1

（1）引出请同学回答上一章的一个问题。

（2）请同学回答一下生活中碰到的与大数据安全相关的事情，引出本章要讲的内容。

（3）教师讲述大数据安全的概念及形成原因。

（4）教师讲述大数据安全的基本分类并介绍不同分类安全问题的形成原因。

（5）请同学回答一下生活中碰到的与大数据隐私相关的场景。

（6）教师讲述大数据隐私问题的发展历程。

（7）教师讲述大数据隐私问题的保护策略。

* 课时2

（1）教师讲授大数据加密技术，首先讲解同态加密的概念及操作步骤。

（2）老师对学生进行现场提问，掌握同学的学习情况，并根据回答确定是否进行再次讲解。问题：同态加密的步骤。

（3）教师讲授可搜索加密的概念及操作步骤。

（4）教授讲授大数据真实性分析认证技术。

（5）教授讲授数据签名的概念、原理、过程及验证过程。

（6）教授讲授基于数据挖掘的认证技术的概念、原理、过程及验证过程。

* 课时3

（1）教师讲授访问控制技术的相关方法。包括基于角色的访问控制、基于属性加密的访问控制以及基于风险的访问控制。

（2）教师讲授数据溯源技术的相关方法。包括反向查询法、基于规则的安全审计、基于统计的安全审计以及基于机器自学习的安全审计。

（3）教师介绍APT攻击检测技术。

（4）教师介绍数据隐藏技术、数据脱敏、数据发布匿名技术等隐私保护技术。

（5）教师介绍大数据在网络安全中的应用。

（6）教师总结全章。

1. **课后作业**

简要分析大数据生命周期中存在哪些安全风险，我们可以使用哪些技术手段实现安全目标？

第八章 大数据在不同行业的应用案例分析

1. **教学目标**
2. 了解大数据在不同行业的应用与前沿；
3. 能够运用大数据相关知识、原理与技术，分析和研究不同行业中的复杂问题。
4. **教学内容**

大数据在不同行业的应用与案例分析。

1. **教学重点与难点**
2. 本章的教学重点是大数据技术在城市计算和金融领域的应用，学生能够运用大数据相关知识、原理和技术对相关领域中的不同复杂问题进行分析；
3. 城市计算是一个交叉学科,是计算机科学以城市为背景,跟城市规划、交通、能源、环境、社会学和经济等学科融合的新兴领域。城市计算是大数据技术在现代城市治理中的典型应用；教学的重点是学生明白城市计算的数据类型、建模方法、城市计算的框架与概念、城市计算的关键技术；
4. 金融科技是基于大数据、云计算、人工智能、区块链等一系列技术创新，全面应用于支付清算、借贷融资、财富管理、零售银行、保险、交易结算等六大金融领域。在课程中，主要讲授大数据技术在金融领域的应用——征信、信贷、反洗钱等。
5. **教学方法与手段**

本章是该课程中前沿知识章节，对本章内容的掌握不能仅仅依靠记忆，更需要学生深刻理解大数据技术、相关行业需求等。因此本章的教学必须通过行业应用的框架和需求出发，通过实例讲解，为学生讲授大数据技术在行业中的应用。

本章以教师讲授（多媒体+板书）为主，师生交互、学生讨论为辅的教学方法与手段进行教学。

1. **教学过程**

* 课时1

（1）请同学们讨论当前城市所面临的问题，注意引导到交通拥堵、城市污染、道路规划等问题。并请同学们提出针对性的解决方案，有没有办法通过大数据技术解决上述问题？

（2）引出城市计算的概念，指出大数据技术在解决城市问题的优势。

（3）教师介绍城市计算的概念、发展情况，通过几个简要实例描绘大数据技术在城市中的应用，着重强调城市计算的概念。

（4）教师讲授城市计算的概念与框架，包括城市数据的类型、特点等，其中时空数据是重点。

（5）教师进一步讲授城市计算的关键技术和应用实例。如空气质量检测、交通拥堵缓解、城市应急系统、城市道路规划等。通过实例将城市计算的建模方式、关键技术穿插讲解。

* 课时2

（1）教师讲授金融科技的相关概念，利用各类科技手段创新传统金融行业所提供的产品和服务，提升效率并有效降低运营成本。

（2）教师讲授大数据技术在金融中的作用，金融行业数据量大，价值密集，大数据技术能够充分挖掘海量数据背后的潜在价值。

（3）教师讲授金融大数据的典型应用案例——征信。

（4）教师讲授金融大数据的典型应用案例——信贷。

（5）教师讲授金融大数据的典型应用案例——反洗钱

1. **课后作业**

简要分析大数据生命周期中存在哪些安全风险，我们可以使用哪些技术手段实现安全目标？