

<<数据挖掘原理与算法>>

图书基本信息

书名：<<数据挖掘原理与算法>>

13位ISBN编号：9787302193500

10位ISBN编号：7302193509

出版时间：2009-6

出版时间：清华大学出版社

作者：毛国君，段立娟 编著

页数：78

字数：134000

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

<<数据挖掘原理与算法>>

前言

《数据挖掘原理与算法》一书出版以来，被许多高校作为本科生或者研究生的教材使用。几年来许多教师给出了很好的建议，因此我们在2007年针对相关问题进行了修订并出版了其第二版。该教材是一本全面介绍数据挖掘和知识发现技术的专业书籍，具有内容系统、知识含量高等特点。可能也正是因为这些特点，作为教材给教师带来了一些授课难点。

特别是，由于教材使用的对象不同，对教材内容进行选择是必需的。

为了让教师更好地使用《数据挖掘原理与算法》一书，减轻教师的负担，我们编写了本教师用书。

《数据挖掘原理与算法（第二版）教师用书》主要从四个部分为教师提供了参考：一、对教材每章的部分习题给出了参考答案；二、介绍各章授课内容重点与课时分配；三、针对不同的授课学生对象给出了课时安排的建议；四、提供了两套样本试卷及其参考答案。

目的是为了帮助教师提高讲课的效率，但不能代替教师的教学研究工作。

特别考虑到教师用书也可能被学生使用，故对教材后面的习题并没有给出全部解答。

整体上说，数据挖掘技术包含概念与过程、原理与方法两个主要部分。

对于有关概念与过程，主要集中在《数据挖掘原理与算法》（第二版）第1章和第2章，不论学生对象如何，教师都应该给予重视，力求全面而直观地进行介绍。

数据挖掘中的原理与方法，分布在《数据挖掘原理与算法》（第二版）的第3—8章，涵盖关联规则、分类、聚类、序列、空间以及Web挖掘等分支。

我们认为，关联规则、分类、聚类是经典内容，不论学生对象如何，教师都应该选择一些典型的理论和算法进行剖析。

对于不同的教学对象，教师可以对第3—5章的内容进行合理选择。

例如，如果准备给本科生开一个只有32课时的课程，那么最起码的要求是在对于关联规则、分类、聚类等基本概念和原理讲述清楚的前提下，能把Apriori、ID3和k-means算法剖析清楚即可。

第6—8章的内容相对比较松散，对于研究生来说，我们认为需要进行选择性地介绍或讨论。

这是因为这些内容属于数据挖掘的较前沿的课题，而且有着很广泛的研究和应用价值，因此对于研究生将来的研究工作可能会有很大的帮助。

<<数据挖掘原理与算法>>

内容概要

《数据挖掘原理与算法》一书出版以来，被许多高校作为本科生或者研究生教材使用，是一本全面介绍数据挖掘和知识发现技术的专业书籍，具有内容系统、知识含量高等特点。

为了让教师更好地使用教材《数据挖掘原理与算法》（第二版），作者又编写了本书。

本书分四个部分：一、对教材每章的部分习题给出了参考答案；二、介绍各章授课内容重点与课时分配；三、针对不同的授课学生对象给出了课时安排的建议；四、提供了两套样本试卷及其参考答案。

本书供使用《数据挖掘原理与算法》一书的教师作参考。

<<数据挖掘原理与算法>>

书籍目录

第一部分 各章习题及部分 参考答案 第1章 绪论 第2章 知识发现过程与应用结构 第3章 关联规则挖掘理论和算法 第4章 分类方法 第5章 聚类方法 第6章 时间序列和序列模式挖掘 第7章 Web挖掘技术 第8章 空间挖掘第二部分 各章 授课重点与课时分配 第1章 绪论 第2章 知识发现过程与应用结构 第3章 关联规则挖掘理论和算法 第4章 分类方法 第5章 聚类方法 第6章 时间序列和序列模式挖掘 第7章 Web挖掘技术 第8章 空间挖掘第三部分 按总学时规划的教学大纲 48学时的教学大纲(本科生) 32学时的教学大纲(本科生) 48学时的教学大纲(研究生)第四部分 样本试卷 样本试卷1(本科生) 样本试卷2(研究生) 样本试卷1(本科生)的参考答案 样本试卷2(研究生)的参考答案

<<数据挖掘原理与算法>>

章节摘录

在关系型数据库的研究和产品提升过程中，人们一直在探索组织大型数据和快速访问的相关技术。高性能关系数据库引擎以及相关的分布式查询、并发控制等技术的使用，已经提升了数据库的应用能力。

在数据的快速访问、集成与抽取等问题的解决上积累了经验。

数据仓库作为一种新型的数据存储和处理手段，被数据库厂商普遍接受并且相关辅助建模和管理工具快速推向市场，成为多数据源集成的一种有效的技术支撑环境。

因此，人们已经具备利用多种方式存储海量数据的能力。

这些丰富多彩的数据存储、管理以及访问技术的发展，为数据挖掘技术的研究和应用提供了丰富的土壤。

计算机芯片技术的发展，使计算机的处理和存储能力日益提高。

随之而来的是硬盘、CPU等关键部件的价格大幅度下降，使得人们收集、存储和处理数据的能力和欲望不断提高。

经过几十年的发展，计算机的体系结构，特别是并行处理技术已经逐渐成熟和普遍应用，并成为支持大型数据处理应用的基础。

计算机性能的提高和先进的体系结构的发展使数据挖掘技术的研究和应用成为可能。

历经了十几年的发展，包括基于统计学、人工智能等在内的理论与技术性成果已经被成功地应用到商业处理和分析中。

这些应用从某种程度上为数据挖掘技术的提出和发展起到了极大地推动作用。

数据挖掘系统的核心模块技术和算法都离不开这些理论和技术的支持。

从某种意义上讲，这些理论本身的发展和应用于数据挖掘提供了有价值的理论和应用积累。

6.数据挖掘技术是一个交叉研究分支，简述影响它产生和发展的主要研究学科或分支及其关系。

参考答案：略。

7.数据（Data）、信息（Information）和知识（Knowledge）是人们认识和利用数据的三个不同阶段，数据挖掘技术是如何把它们有机的结合在一起的？

参考答案：从数据、信息和知识三个层面上看，数据是最原始的未经组织和处理的信息源。

信息或称有效信息是指对人们在某些方面有价值的东西。

知识是二种现实世界信息的抽象和浓缩，是一种概念、规则、模式和规律等。

数据挖掘技术通过对原始数据进行微观、中观乃至宏观的统计、分析、综合和推理，发现数据间的关联性、未来趋势以及一般性的概括知识等，转变成可以用来指导人们某些高级商务活动的有用信息。

8.从数据挖掘研究角度看，如何理解数据、信息和知识的不同和联系。

参考答案：略。

9.简述数据挖掘技术将来的发展趋势。

参考答案：对于数据挖掘技术的发展趋势，应该分两方面辩证的理解。

（1）数据挖掘技术已经存在相当大市场，将成为对工业产生重要影响的关键技术之一。

同时，并行计算机体系结构研究和KDD也被列入今后5年内公司应该投资的10个新技术领域之一。这些资料都表明，数据挖掘技术在将来有很大的发展潜力及空间。

<<数据挖掘原理与算法>>

编辑推荐

本书供使用《数据挖掘原理与算法》一书的教师作参考。

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>