

<<数据仓库与数据挖掘>>

图书基本信息

书名：<<数据仓库与数据挖掘>>

13位ISBN编号：9787301143131

10位ISBN编号：7301143133

出版时间：2008-11

出版时间：北京大学出版社

作者：廖开际 编

页数：251

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>

## &lt;&lt;数据仓库与数据挖掘&gt;&gt;

## 前言

随着电子商务与信息管理技术及应用在我国和全球的迅速发展,政府、行业和企业对电子商务与信息管理的重视程度不断提高,我国高校电子商务与信息管理人才培养的任务也不断加重。作为一个新兴的跨学科领域的专业,电子商务与信息管理的教育在快速发展的同时还存在着许多值得我们思考和改进的问题。

特别是开办电子商务专业和信息管理专业的学校学科背景不同,有文科的、理工科的、经管类学科等,使得不同学校对核心课程的设置差异很大;另外,近年来有关电子商务与信息管理方面的教材出版的数量虽然不少,但适合于财经管理类知识背景本科生的电子商务系列与信息管理系列教材一直缺乏,而在开办电子商务和信息管理本科专业的高校中,财经管理类的高校占的比重很大。

为此北京大学出版社于2006年11月在北京召开了《21世纪全国应用型本科财经管理系列实用规划教材》研讨会暨组稿会,会上出版社的领导和编辑通过对国内经管类学科背景的多所大学电子商务与信息管理系列教材实际情况的调研,在与众多专家学者讨论的基础上,决定成立电子商务与信息管理系列丛书专家编审委员会,组织编写和出版一套面向经管类学科背景的电子商务与信息管理专业的应用型系列教材,暨《21世纪全国应用型本科电子商务与信息管理系列实用规划教材》。

本系列教材的特点在于,按照高等学校电子商务专业与信息管理专业对本科教学的基本要求,参考教育部高等学校电子商务专业与信息管理专业的课程体系和知识体系,定位于实用型人才培养。

本系列教材还体现了教育思想和教育观念的转变,依据教学内容、教学方法和教学手段的现状和趋势进行了精心策划,系统、全面地研究普通高校教学改革、教材建设的需求,优先开发其中教学急需、改革方案明确、适用范围较广的教材。

此次教材建设的内容、架构重点考虑了以下几个要素。

(1) 关注电子商务与信息管理发展的大背景,拓宽经济管理理论基础、强调计算机应用与网络技术应用技能和专业知识,着眼于增强教学内容的联系实际和应用性,突出创造能力和创新意识。

(2) 尽可能符合学校、学科的课程设置要求。

以高等教育的培养目标为依据,注重教材的科学性、实用性和通用性,尽量满足同类专业院校的需求。

(3) 集中了在电子商务专业与信息管理专业教学方面具有丰富经验的许多教师和研究人员的宝贵意见,准确定位教材在人才培养过程中的地位和作用。

面向就业,突出应用。

f4) 进行了合理选材和编排。

教材内容很好地处理了传统内容与现代内容的关系,补充了大量新知识、新技术和新成果。

根据教学内容、学时、教学大纲的要求,突出了重点和难点。

## <<数据仓库与数据挖掘>>

### 内容概要

《数据仓库与数据挖掘》比较系统地介绍数据仓库与数据挖掘的理论体系和应用。

《数据仓库与数据挖掘》总的指导思想是在掌握基本知识和基本理论的基础上，强调实际应用能力的培养。

全书力求深入浅出，通过通俗的语言及案例分析，介绍数据仓库及数据挖掘的基本概念及相关理论与方法。

从数据仓库的定义、结构、设计、构建方法及联机分析处理应用等方面对数据仓库进行较为详细的介绍；从数据挖掘的定义、数据预处理、数据挖掘中的常用算法等方面对数据挖掘的基本知识和算法等理论进行介绍。

《数据仓库与数据挖掘》强调数据仓库和数据挖掘工具的应用，重点介绍SQL Server 2005数据仓库和数据挖掘工具的应用。

附录A详细介绍一个简易的数据挖掘工具——Weka，该工具可作为读者学习数据挖掘时的实验工具。

《数据仓库与数据挖掘》可作为普通高等学校电子商务、信息管理、计算机应用及其他相关专业的本科教材，也可作为经贸、管理类专业的研究生教材，以及各类管理人员的培训与自学用书。

## &lt;&lt;数据仓库与数据挖掘&gt;&gt;

## 书籍目录

第1章 企业数据资源管理1.1 数据资源的概念1.1.1 企业资源1.1.2 数据资源1.1.3 数据资源管理及其发展历程1.2 数据资源管理的意义1.2.1 信息系统进入成熟阶段的重要标志1.2.2 解决企业内部数据不一致问题的根本途径1.2.3 数据资源的管理和应用是取得竞争优势的关键1.3 信息资源管理的相关技术1.3.1 数据资源管理的技术框架1.3.2 技术框架中的构成要素1.3.3 技术框架中各部分的关联1.4 企业通过数据仓库与数据挖掘获得竞争优势本章小结思考与练习第2章 数据仓库的概念与结构2.1 数据仓库的概念2.1.1 数据仓库的定义2.1.2 数据仓库的特征2.1.3 数据集市2.2 数据仓库系统2.2.1 数据源2.2.2 数据仓库管理层2.2.3 数据仓库工具集2.3 数据仓库中的数据组织2.3.1 粒度的概念2.3.2 面向主题的数据组织2.3.3 数据分割2.3.4 元数据的管理本章小结思考与练习第3章 数据仓库的设计与开发3.1 数据仓库的开发过程及特点3.1.1 数据仓库开发生命周期3.1.2 数据仓库开发的特点3.1.3 数据仓库设计的主要内容3.2 数据模型设计3.2.1 概念模型设计3.2.2 逻辑模型设计3.2.3 物理模型设计3.3 数据仓库的粒度设计3.3.1 设计步骤3.3.2 设计原则3.4 创建数据仓库的基本步骤3.4.1 建立运营环境文档3.4.2 选择数据仓库的实现技术3.4.3 设计数据仓库模型3.4.4 创建数据准备区3.4.5 创建数据仓库数据库3.4.6 从操作型系统中抽取数据3.4.7 清理和转换数据3.4.8 将数据装入数据仓库数据库3.4.9 准备显示信息3.4.10 将数据分发到数据集市本章小结思考与练习第4章 联机分析处理4.1 OLAP的基本概念4.1.1 OLAP的发展背景4.1.2 联机分析处理是数据仓库系统的一个应用4.2 OLAP与多维分析4.2.1 OLAP的一些基本概念4.2.2 理解数据立方4.2.3 OLAP的基本分析操作4.3 OLAP的分类4.3.1 ROLAP4.3.2 MOLAP4.3.3 HOLAP4.4 OLAP的特性与不足4.4.1 OLAP的特性4.4.2 OLAP的不足4.5 SQL Server 2005统一维度模型4.5.1 结构4.5.2 优点本章小结思考与练习第5章 数据挖掘概述5.1 数据挖掘技术的由来5.1.1 信息爆炸但知识贫乏5.1.2 支持数据挖掘技术的基础5.1.3 从商业数据到商业信息的进化5.1.4 数据挖掘逐渐演变的过程5.2 数据挖掘的定义5.2.1 技术角度的定义5.2.2 商业角度的定义5.2.3 数据挖掘与传统分析方法的区别5.2.4 数据挖掘和数据仓库5.2.5 数据挖掘和OLAP5.2.6 数据挖掘、机器学习和统计5.3 数据挖掘发现的知识类型5.3.1 广义知识5.3.2 关联知识5.3.3 分类知识5.3.4 预测知识5.3.5 偏差知识5.4 数据挖掘流程5.4.1 知识发现过程5.4.2 数据挖掘对象5.4.3 数据挖掘任务5.4.4 数据挖掘分类5.4.5 数据预处理5.5 数据挖掘的方法和技术5.5.1 信息论方法5.5.2 集合论方法5.5.3 神经网络方法5.5.4 遗传算法5.5.5 模糊数学5.5.6 公式发现5.5.7 可视化技术5.5.8 知识表示本章小结思考与练习第6章 数据预处理6.1 数据预处理的目的是方法6.1.1 原始数据中存在的问题6.1.2 数据预处理的常用方法6.2 数据清理6.2.1 处理空缺值6.2.2 噪声数据的处理6.3 数据集成6.3.1 模式匹配6.3.2 数据冗余6.3.3 数据冲突6.4 数据变换6.5 数据归约6.5.1 数据立方体聚集6.5.2 维归约6.5.3 数据压缩6.5.4 数值归约6.5.5 离散化和概念分层本章小结思考与练习第7章 数据挖掘中的常用算法7.1 Apriori算法7.1.1 基本原理7.1.2 Apriori算法的基本思想与分析7.1.3 从频繁项集产生关联规则7.2 决策树算法7.2.1 信息论的基本原理7.2.2 ID3算法7.2.3 树剪枝7.2.4 由决策树提取分类规则7.3 神经网络算法7.3.1 神经网络的基本原理7.3.2 反向传播模型7.3.3 定义神经网络拓扑结构7.3.4 神经网络的工作过程7.4 聚类分析7.4.1 聚类分析的概念7.4.2 聚类分析中的数据类型7.4.3 几种主要的聚类分析方法7.4.4 K means聚类分析算法本章小结思考与练习第8章 SQL Server数据仓库与数据挖掘工具及其应用8.1 SQL Server 2005的功能构架8.2 SQL Server数据仓库设计与数据挖掘准备8.2.1 SQL Server数据仓库创建思路8.2.2 SQL Server数据挖掘过程8.2.3 案例数据准备8.3 SQL Server集成服务8.3.1 SQL Server集成服务的作用8.3.2 控制流8.3.3 数据流8.3.4 设计和使用ETL8.4 SQL Server分析服务8.4.1 创建Analysis Services项目8.4.2 定义数据源8.4.3 定义数据源视图8.4.4 用Analysis Services创建维与多维数据集8.4.5 部署Analysis Services项目8.5 SQL Server中的数据挖掘工具与应用8.6 SQL Server报表服务8.6.1 创建报表8.6.2 使用报表本章小结思考与练习附录A 一个简易的数据挖掘工具——Weka参考文献

## 章节摘录

第1章 企业数据资源管理 1.2 数据资源管理的意义 数据资源管理的意义可从以下3个方面来理解。

(1) 从信息系统的发展过程来看,对数据资源进行有效的管理是信息系统进入成熟阶段的重要标志。

(2) 数据资源管理是解决企业内部由于数据重复而导致的各种问题的根本途径。

(3) 数据资源是企业取得竞争优势的关键。

1.2.1 信息系统进入成熟阶段的重要标志 企业的信息系统的成长过程分为5个阶段:初级、普及、整理、集成、成熟。

在初级阶段,计算机刚进入企业,只作为办公设备使用,应用非常少,通常用来完成一些报表统计工作,甚至大多数时候被当做打字机使用。

随着企业对计算机应用认识的深入,人们体会到计算机应用的价值,开始学习、使用和维护计算机。

在普及阶段,计算机应用在一些部门见到成效,从最初的一些应用部门向其他部门蔓延,大量的人工数据处理转向计算机处理,人们对计算机的热情增加,需求增长。

在整理阶段,由于人们对计算机信息处理需求的增长,造成财务支出的大幅度上涨,企业领导不得不对之进行控制,注重采用成本/效益去分析应用开发。

并针对各项已开发的应用项目之间的不协调和数据冗余等问题进行统一规划。

这一阶段的效益可能比第二阶段还要低。

在集成阶段,企业高层领导意识到信息战略的重要,信息成为企业的重要资源,企业的信息化建设也真正进入到数据处理阶段。

在这一阶段中,工作的重点是对数据资源进行管理和控制,包括数据的质量控制、数据的有效利用和数据价值的提升。

企业开始选定统一的数据库平台、数据管理体系和信息管理平台,以统一数据的管理和使用,各部门、各系统基本实现资源整合、信息共享。

在成熟阶段,信息系统已经可以满足企业各个层次的需求,从简单的事务处理到支持高效管理的决策。

工作的重点是建立以数据资源为基础的系统计划和战略计划,将信息系统作为取得竞争优势的有力手段。

.....



## <<数据仓库与数据挖掘>>

### 编辑推荐

丛书特点： 1.突出创造能力和创新意识。

关注专业背景，拓宽理论基础、强调计算机应用与网络技术应用技能和专业知识，着眼于增强教学内容的实际和应用性。

2.符合各学校专业课程设置要求。

以高等教育的培养目标为依据，注重教材的科学性、实用性和通用性，准确定位教材在人才培养过程中的地位和作用，满足各院校教学需求。

3.面向就业，突出应用。

作者多为在电子商务与信息管理专业教学方面具有丰富经验的一线教师和研究人員，准确把握就业市场动向，注重培养学生实际操作能力。

4.合理选材和编排。

传统内容与现代内容合理融合，补充了大量新知识、新技术和新成果；遵循最新准则或规范，根据教学内容、学时、教学大纲的要求，突出重点和难点。

5.侧重案例教学。

对大量当前最新典型案例进行分析讲解，理论联系实际，通俗易懂。

《数据仓库与数据挖掘》共分为8章和1个附录。

第1章介绍企业数据资源管理，旨在说明数据仓库与数据挖掘是企业数据资源管理的高级阶段，也是必然趋势。

同时也介绍了数据仓库与数据挖掘能为企业做什么。

第2章介绍数据仓库的概念与结构，读者在学完本章后可以明白什么是数据仓库，数据仓库是怎样构成的。

第3章介绍数据仓库的设计与开发，在学完本章后，读者可以根据需求设计自己的数据仓库，并逐步地建立自己的数据仓库（或数据集市）。

第4章介绍数据仓库的一个重要应用——联机分析处理，读者可以理解多维数据分析的方法。

第5章介绍数据挖掘概述，在学完本章后，读者应能明白数据挖掘能做什么，有些什么方法。

第6章介绍数据挖掘中的一个重要的也是工作量最大的一个环节——数据预处理，在学完本章后，读者应明白为什么要进行预处理，预处理的各种方法，并学会怎样进行数据的预处理。

第7章介绍数据挖掘中的常用算法，读者可以理解这些算法的基本思想，掌握典型算法的应用。

第8章结合一个实际案例介绍SQL Server 2005数据仓库与数据挖掘工具及其应用。

建议读者在本章学习的基础上，自己深入钻研SQL Server 2005的数据仓库与数据挖掘工具。

《数据仓库与数据挖掘》适合作为非计算机专业的教材及自学读物。

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>