



普通高等教育“十一五”国家级规划教材
高等学校信息管理示范教材

数据仓库与 数据挖掘技术(第2版)

■ 陈京民 编著

<http://www.phei.com.cn>



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

普通高等教育“十一五”国家级规划教材
高等学校信息管理示范教材

数据仓库与数据挖掘技术

(第2版)

陈京民 编著

电子工业出版社
Publishing House of Electronics Industry
北京·BEIJING

内 容 简 介

本书介绍了以数据仓库、数据挖掘和联机分析为核心技术的商务智能的基本概念、基本原理、开发方法、开发工具、应用领域和管理方法等内容。

全书共分为 10 章,包括商务智能基本概念、数据仓库开发模型、数据仓库的开发应用过程、OLAP 技术、传统数据挖掘技术、现代数据挖掘技术与发展、电子商务中的商务智能、商务智能应用与管理,以及商务智能系统的开发应用实例等内容。每章后还附有可供读者自我测试的习题,以帮助读者对全书内容的理解。本书内容翔实,结构清晰,应用性强,既可以作为高等学校信息管理与信息系统专业、计算机应用专业、自动控制专业以及相关专业的本科生和研究生的教材,也可以作为企业、事业单位从事商务智能系统开发应用工作人员的参考用书。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究。

图书在版编目(CIP)数据

数据仓库与数据挖掘技术/陈京民编著. —2 版. —北京:电子工业出版社,2007. 11
普通高等教育“十一五”国家级规划教材 高等学校信息管理示范教材
ISBN 978-7-121-05305-4

I. 数… II. 陈… III. ① 数据库系统 ② 数据采集 IV. TP311.13 TP274

中国版本图书馆 CIP 数据核字(2007)第 172463 号

策划编辑:刘宪兰

责任编辑:刘宪兰

印 刷:北京市海淀区四季青印刷厂

装 订:涿州市桃园装订有限公司

出版发行:电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本:787×1092 1/16 印张:24.25 字数:572 千字

印 次:2007 年 11 月第 1 次印刷

印 数:5000 册 定价:32.00 元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010)88254888。

质量投诉请发邮件至 zls@phei.com.cn,盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线:(010)88258888。

第1版序

20世纪70年代,当强大的信息化巨潮还蕴藏在大洋深处,我们的陆地只有一阵微风吹来之时,有识之士们就开始推动信息化专业人才的培养计划,为迎接即将到来的信息化巨潮扩军备战。他们一方面推动着信息技术的普及;一方面根据不同领域的需求,从不同的角度创办了不同类型的信息化专业,这就是管理信息系统专业、经济信息管理专业、科技信息管理专业、医学信息管理专业、林业信息管理专业、农业信息管理专业……实际上,这些专业培养目标可以概括为:为各行业、各部门培养以CIO为目标的信息化专门人才。从这一点上看,这些专业的课程设置应当具有相当大的共同性。1996年,出于多种考虑,教育部将这些专业合并为一个——信息管理与信息系统专业。

以CIO为目标的信息化专门人才是一类管理人才。但是他们所管理的主要对象是信息。这样的知识需求,将信息管理与信息系统专业定位于管理学科,与信息学、经济学、法学等学科交叉。这样的学科特点,给课程建设和教材建设带来不少困难。近30年来,尽管我们与许多同行已经进行了不懈的努力,把信息管理与信息系统专业的课程建设和教材建设向前推进了一大步,但是仍然不尽如人意,许多课程和教材还没有体现信息管理专业的特色和需要。在多次有关的研讨会上,大家一致呼吁编写一套真正体现信息管理与信息系统专业特色的教材。

新编和出版一套专业教材是要冒风险的,而编写和出版一套以瞬息万变的信息和信息技术为管理对象的专业教材就要冒更大的风险。国内信息业界著名的出版商——电子工业出版社,以超人的胆略愿意同我们一道承担这一风险,组织编写出版一套新的信息管理与信息系统专业核心教材。这套教材冠以“新编”二字,是试图在其体系上能比已有教材更体现信息专业的特色,同时在内容上更能反映最新信息技术的进步以及最新信息管理思想和方法。目前,国内开设信息管理与信息系统专业的高等院校已经超过200所。这样一个数字一方面表明信息化已经深入人心,信息化队伍的规模正在急速扩大,信息化队伍的素质正在不断提高;另一方面,也给我们增加了巨大的压力,使我们深感责任重大。好在国内本领域的三位知名学者——黄梯云、陈禹、马费成以及其他一批著名专家和后起之秀愿意与我们共担风险,鼓舞了我们挑起这副重担的勇气。同时,我们也把这套教材的不断精化寄希望于广大的同仁,愿我们把这套教材越改越好,永改永新。

编委会
2002年5月

第2版序

管理作为有效实现目标的社会活动,自古有之。古代的中国人、巴比伦人、苏美尔人、古埃及人、希伯来人、古希腊人、古罗马人,都创立了许多管理思想。但是现代西方管理的基本思想是与近代大工业生产及科学技术的发展紧密联系在一起,例如亚当·斯密的管理思想是与第一次工业革命联系在一起的。从此开始,管理思想不断发展,如泰罗、吉尔布雷斯、甘特、福特等人的科学管理,法约尔、韦伯等人的组织管理;梅奥等人的行为管理等,马斯洛的需求层次理论,赫茨伯格的双因素理论等。随着计算机的出现,人类处理信息的能力得到极大的提高,也同时认识到信息资源的能动作用,管理的核心随之转移到了信息之上——信息管理应运而生。随着世界性的信息化浪潮的迅速推进,信息管理扩展到了各行各业,又形成电子商务、电子政务、企业信息化、医院信息化……多个子领域,迅速形成一个庞大而独立的专业领域和学科范畴,仅2005年的初步统计,我国信息管理本科专业已经有500多个布点。这套教材就是为满足这样的教学要求,于2001年以“新编信息管理与信息系统核心教材”为名开始组织编写的。

从目前看,信息管理专业大致可以分为两个大的方向:信息系统建设与管理 and 信息资源建设与管理。在具体教学中各个学校大都采取了以其中一个方向为主兼顾另一个方向的做法。所以,我们从一开始,就把这套书定位在二者兼顾上。

教学是一个严肃的过程,教材的质量是教学的生命线。为了保证这套教材的质量,每本书的作者都是在充分调研的基础上确定的,在编写过程中主编、作者和编辑反复沟通。与此同时,我们还聘请了这个领域有代表性的知名学者——黄梯云、陈禹、马费成作为顾问,并聘请有关专家参加编辑委员会的工作,层层把关。在大家的努力下,这套书的质量得到了社会的肯定,在2006年公布的国家“十一五”规划教材中,这套书的大部分都列入其中。这一结果鼓舞我们把这套书编写得更好。我们也把这个结果作为一个新的起点,并按照大家的建议,把这套教材更名为“信息管理示范教材”。

“示范”就是抛砖引玉,希望通过我们的努力,把信息管理专业教材的质量提高到一个新的高度。同时,也希望广大读者提出批评、建议和指导。

编委会
2007年6月

第1版前言

信息技术的迅速发展已将我们从简单的批处理、联机事务处理的信息处理时代,带入了联机分析处理、数据仓库和数据挖掘的信息分析时代。数据仓库在短短的几年内已经从一种单纯的理论研究发展成信息管理与信息系统开发领域中一种实用性极强的技术。这一发展具有其内在的动力和外在的推动。由于大多数企业在早期的信息化进程中构建了比较完善的信息处理系统——联机事务处理系统,这些系统为企业的业务快速、准确处理提供基本条件,同时为企业积累大量的、有价值的业务信息。这些处理只能支持企业的日常业务工作,而对企业的经营管理决策却无法提供支持。许多企业的经营管理人员在日趋严重的市场竞争压力下,开始着手建立自己的数据存储——数据集市,用于经营管理决策,以应对日益严酷的市场竞争压力。这些因素最终促进了数据仓库技术的发展。

数据仓库的建立不仅需要各种建设工具,而且还需相应的数据支持。数据仓库的建设必须基于比较完善的信息化构架,只有在一定的信息化基础上,才能进行数据仓库的建设。数据仓库的建设还是企业经营管理决策与信息化结合的过程,只有依照企业的管理决策实际情况,才能建设一个支持企业管理决策的数据仓库。数据仓库的建设还是各种先进的信息处理技术与企业管理决策结合的过程,只有将 OLAP 技术、数据挖掘技术与数据仓库中的庞大数据相结合,与企业先进的管理决策方法相结合,才能使数据仓库在企业的经营管理决策中发挥巨大的作用。数据仓库建设的成功不仅取决于技术人员对数据仓库开发方法与开发工具的熟练应用,更重要取决于数据仓库能否得到熟练应用。可以毫不夸张地说,数据仓库的成功关键在于用户的应用情况,而不是数据仓库开发技术的熟练应用。因此,本书在集中介绍数据仓库的开发模型与开发方法后,还用相当的篇幅介绍数据仓库的管理与应用。

本书还介绍了大量的数据仓库应用情况与应用案例,使读者了解如何利用数据仓库来降低企业的运营成本、建立更好的客户关系管理、提高产品的质量等。本书还介绍数据仓库开发应用的生命周期,数据仓库的整个开发过程从规划分析到设计实施、终结于应用管理,使读者了解数据仓库开发应用的完整周期,以及如何处理在不同阶段中所遇到的问题。为使读者通过实际的数据仓库开发利用,加深对数据仓库与数据挖掘的了解,本书还介绍一些数据仓库与数据挖掘的工具,使读者通过这些工具的实际应用,进一步加强对数据仓库与数据挖掘技术的了解。

全书共分 13 章。第 1 章主要介绍数据仓库的产生背景、发展、用户类型、总体结构与使用技术;第 2 章介绍数据仓库的开发工具——Oracle 9i 在数据仓库开发中的应用;第 3 章介绍另一种数据仓库开发工具——SQL Server 2000 在数据仓库开发应用中的作用;第 4 章介绍数据仓库应用的前台开发工具——Delphi 6.0 在数据仓库开发中的应用;第 5 章从理论上介绍数据仓库的开发模型——概念模型、逻辑模型、物理模型、元数据模型和数据粒度模型;第 6 章叙述数据仓库开发应用的完整周期,涉及数据仓库开发规划、需求分析、设计、实施、使用及支持等;第 7 章阐述数据仓库的实际开发过程与运行过程的技术支持;第 8 章介绍联机分析技术(OLAP)的基本概念、结构、实施以及 OLAP 工具评价标准;

第9章阐述数据挖掘技术的发展过程,数据挖掘工具及其应用;第10章详细介绍统计分析类数据挖掘技术、工具、应用及其应用中的问题;第11章进一步介绍知识类数据挖掘技术与工具;第12章介绍文本挖掘、Web挖掘等其他类型的数据挖掘工具与应用;第13章说明数据仓库的实际应用,应用中的问题和数据仓库开发应用中的管理问题。

本书第1,2,3,4,8章由陈京民编写;第5,13章由杜冬军编写;第6,7章由杜冬军、陈京民编写;第9章由俞强编写;第10,11章由俞强、陈京民编写;第12章由朱惠云编写。全书最后由陈京民修改、统稿。

在本书的编著过程中,张基温教授提供了大力支持与帮助,且对本书提出了非常宝贵的评审与修改意见,在此表示衷心的感谢。

感谢刘宪兰在本书出版工作中所给予的大力支持与帮助。

由于编者水平有限,书中难免存在不当之处,而数据仓库与数据挖掘又正处于日新月异的发展过程之中,恳切希望各位读者批评指正。

电子邮件地址:cjm20020101@sina.com.cn

作者

2002年5月

第2版前言

本书第1版出版以来,得到了广大读者的大力支持与肯定,在此表示衷心的感谢。随着数据仓库与数据挖掘技术的不断成熟与普及,在数据仓库与数据挖掘技术基础上所构成的商务智能理念越来越得到信息管理领域专家、学者的认可,急需将它们补充和充实到本书中。同时,在本书的使用过程中,发现某些概念或问题可以有更好阐述的方法和论述角度,为此借本书被评选为普通高等教育“十一五”国家级规划教材重新修订出版之际,对全书文字叙述进行了修订,使其表达更准确、流畅;对全书内容从新的角度进行了阐述,使修改后的内容更易理解;调整了部分章节结构,使全书安排更加清晰合理;新增了部分内容以反映一些最新的概念、平台工具与开发技术,使读者能够在阅读之余动手设计一个商务智能应用系统。

第2版在第1版写作思路的基础上,对商务智能、数据仓库和数据挖掘技术等知识进行了深入探讨,构造了全书新的体系结构,并深入浅出地阐述了数据仓库与数据挖掘技术,使全书理论与技术相结合、设计与应用相结合。全书主要修订部分如下:

第1章“商务智能基本概念”在原书的基础上,合并了原书第9章部分内容和增加了商务智能概念。

第2章“数据仓库开发模型”在原书第5章的基础上修改而成。

第3章“数据仓库开发应用过程”由原书第6、7章合并而成。

第4章“OLAP技术”为原书第8章,保持了原章节的体系结构。

第5章“传统数据挖掘技术”在原书第10章的基础上,对其内容稍作修改。

第6章“现代数据挖掘技术与发展”由原书第11章和12章合并而成,除对一些数据挖掘内容进行精简外,增加了一些数据挖掘的算法。

第7章“电子商务中的商务智能”为新增内容,主要对电子商务数据中的商务智能应用进行了讨论,其中主要涉及Web数据仓库的开发与Web数据的挖掘。

第8章“商务智能应用与管理”由原书第13章与第9章的4节组成,并增加了商务智能在企业 and 行业中的应用等内容,使商务智能的应用与管理内容更加完善。

第9章“商务智能开发实例”是在原书第3章的基础上修订而成,主要包含了基于商务智能的超市系统数据仓库架构的规划与分析、SQL Server 2005的数据仓库创建工具,使读者能够亲身体会到商务智能的实际开发乐趣。

第10章“商务智能应用实例”是在原书第3章的基础上修订而成,内容主要包括应用SQL Server 2005工具开发超市系统数据仓库的实际案例,增加了对商务智能客户端界面的设计案例,使读者能够了解如何实现商务智能的实际应用。

本书主要由陈京民编著并最后统稿。参加本书第1章编写工作的还有祁沁红,参加本书第2章编写工作的还有吴爱珠,参加本书第6章编写工作的还有朱慧云,参加本书第8章编写工作的还有管向梅,参加本书第9、10章编写工作的还有韩松、张振。

电子工业出版社的刘宪兰对本书的编写给予了大力支持,作者在此对她表示衷心的感谢。

限于水平,本书难免有欠妥与错误之处,恳请读者及专家批评指正。敬请通过 cjm@hhu.edu.cn 与作者联系,不胜感谢。

作者

2007年8月

目 录

第 1 章 商务智能基本概念	(1)
1.1 商务智能的基本概念	(2)
1.1.1 商务智能的定义	(2)
1.1.2 商务智能的发展与应用	(3)
1.1.3 商务智能的体系结构	(5)
1.2 数据仓库的发展与展望	(6)
1.2.1 从传统数据库到数据仓库	(6)
1.2.2 数据仓库的定义与基本特性	(9)
1.2.3 数据仓库的发展	(12)
1.3 数据仓库的体系结构	(14)
1.3.1 数据仓库的概念结构	(14)
1.3.2 虚拟数据仓库结构	(14)
1.3.3 数据集市结构	(15)
1.3.4 单一数据仓库结构	(15)
1.3.5 分布式数据仓库结构	(16)
1.4 数据仓库的参照结构	(17)
1.4.1 数据仓库基本功能层	(17)
1.4.2 数据仓库的管理层	(24)
1.4.3 数据仓库的环境支持层	(26)
1.5 数据挖掘技术概述	(28)
1.5.1 数据挖掘的发展	(28)
1.5.2 数据挖掘的定义	(29)
1.6 数据挖掘技术与工具	(31)
1.6.1 常用数据挖掘技术	(31)
1.6.2 常用数据挖掘工具	(33)
1.6.3 数据挖掘工具的评价标准	(35)
1.6.4 常用数据挖掘工具的选择	(36)
1.7 数据挖掘的应用	(37)
1.7.1 数据挖掘与数据仓库	(37)
1.7.2 数据挖掘过程	(38)
1.7.3 数据挖掘的用户	(42)
本章小结	(43)
自我测试题	(43)

第2章 数据仓库开发模型	(47)
2.1 数据仓库开发模型	(48)
2.2 数据仓库概念模型	(49)
2.2.1 概念模型	(49)
2.2.2 规范的数据模型	(52)
2.2.3 星形模型	(53)
2.2.4 雪花模型	(55)
2.3 数据仓库的逻辑模型	(55)
2.3.1 事实表模型设计	(57)
2.3.2 维模型设计	(59)
2.4 数据仓库的物理模型	(59)
2.4.1 数据仓库物理模型的存储结构	(59)
2.4.2 数据仓库物理模型的索引构建	(60)
2.4.3 数据仓库物理模型的优化问题	(63)
2.5 数据仓库的元数据模型	(65)
2.5.1 元数据的类型与组成	(65)
2.5.2 元数据在数据仓库中的作用	(66)
2.5.3 元数据的收集	(70)
2.6 数据仓库的粒度和聚集模型	(71)
2.6.1 数据粒度的划分	(72)
2.6.2 确定粒度的级别	(73)
2.6.3 数据仓库聚集模型的确定	(73)
2.6.4 聚集模型的处理	(74)
2.6.5 聚集模型的管理	(75)
本章小结	(75)
自我测试题	(76)
第3章 数据仓库的开发应用过程	(79)
3.1 数据仓库开发应用的特点	(80)
3.1.1 数据仓库开发应用的阶段性	(80)
3.1.2 数据仓库的螺旋式开发方法	(81)
3.1.3 数据仓库的开发特点	(82)
3.2 数据仓库的规划	(83)
3.2.1 选择数据仓库实现策略	(83)
3.2.2 确定数据仓库的开发目标和实现范围	(84)
3.2.3 选择数据仓库的体系结构	(85)
3.2.4 建立数据仓库使用方案和项目规划预算	(87)
3.3 数据仓库的概念模型设计	(88)
3.3.1 概念模型的需求调查	(88)
3.3.2 概念模型的定义	(89)
3.3.3 概念模型的分析	(93)

3.3.4	概念模型的设计	(94)
3.3.5	概念模型文档与评审	(94)
3.4	数据仓库的逻辑模型设计	(97)
3.4.1	分析主题域	(97)
3.4.2	粒度层次和聚集的确定	(98)
3.4.3	确定数据分割策略	(98)
3.4.4	关系模型的定义	(99)
3.4.5	数据仓库的实体定义	(100)
3.4.6	数据仓库的数据抽取模型	(101)
3.4.7	数据仓库元数据模型的建立与应用	(106)
3.4.8	逻辑模型的评审	(107)
3.5	数据仓库的物理模型设计	(107)
3.5.1	数据仓库设计的规范	(108)
3.5.2	确定数据结构类型	(108)
3.5.3	数据仓库索引的创建	(109)
3.5.4	确定数据存放位置	(111)
3.5.5	确定存储分配	(111)
3.5.6	数据仓库物理模型的评审	(112)
3.6	数据仓库的实施	(113)
3.6.1	数据仓库与业务处理系统接口的设计	(113)
3.6.2	数据仓库的创建	(114)
3.6.3	数据仓库的数据加载、复制与发行	(114)
3.6.4	数据仓库的中间件设计	(115)
3.6.5	数据仓库的测试	(116)
3.7	数据仓库的应用、支持和增强	(117)
3.7.1	数据仓库的用户培训及支持	(117)
3.7.2	数据仓库的使用方式	(118)
3.7.3	数据仓库使用中的数据刷新	(119)
3.7.4	数据仓库的增强	(120)
	本章小结	(121)
	自我测试题	(121)
第4章	OLAP 技术	(125)
4.1	OLAP 技术概念	(126)
4.1.1	OLAP 的发展	(126)
4.1.2	OLAP 的特性	(126)
4.2	OLAP 与多维分析	(127)
4.2.1	多维基本概念	(127)
4.2.2	多维分析	(130)
4.2.3	维的层次关系	(132)
4.2.4	维的类关系	(132)

4.2.5	OLAP 与数据仓库的关系	(133)
4.3	OLAP 的实施	(134)
4.4	多维 OLAP 与关系 OLAP	(135)
4.4.1	多维数据库	(135)
4.4.2	多维数据库的数据存储	(137)
4.4.3	多维数据库与数据仓库	(137)
4.4.4	MOLAP 的创建与功能	(139)
4.4.5	ROLAP 实现的 3 个规则	(139)
4.4.6	ROLAP 的多维表示方法	(140)
4.4.7	ROLAP 的创建与功能	(142)
4.5	OLAP 技术的评价	(143)
4.5.1	MOLAP 与 ROLAP 的比较	(143)
4.5.2	OLAP 的衡量标准	(145)
4.5.3	OLAP 服务器和工具的评价标准	(147)
	本章小结	(149)
	自我测试题	(149)
第 5 章	传统数据挖掘技术	(153)
5.1	传统的统计分析类数据挖掘技术	(154)
5.1.1	统计与统计类数据挖掘技术	(154)
5.1.2	数据的聚集与度量技术	(155)
5.1.3	柱状图数据挖掘技术	(155)
5.1.4	线性回归数据挖掘技术	(156)
5.1.5	非线性回归数据挖掘技术	(158)
5.1.6	聚类数据挖掘技术	(159)
5.1.7	最近邻数据挖掘技术	(165)
5.2	统计分析类工具	(166)
5.2.1	统计类数据挖掘工具	(166)
5.2.2	统计类数据挖掘的商业分析	(167)
5.2.3	统计类数据挖掘工具的功能	(168)
5.2.4	统计类数据挖掘工具—SPSS	(169)
5.3	统计分析类工具的应用	(172)
5.3.1	趋势分析	(172)
5.3.2	时序分析	(173)
5.3.3	周期分析	(173)
5.4	统计分析类工具应用的问题	(174)
5.4.1	统计类数据挖掘的预处理问题	(174)
5.4.2	统计分析应遵循的基本原则	(176)
5.4.3	统计分析的步骤	(177)
5.4.4	统计类数据挖掘的性能问题	(178)
	本章小结	(178)

自我测试题	(179)
第 6 章 现代数据挖掘技术与发展	(183)
6.1 知识挖掘系统的体系结构	(184)
6.1.1 知识发现技术的定义	(184)
6.1.2 知识发现系统的结构	(185)
6.2 现代挖掘技术及应用	(187)
6.2.1 规则型挖掘技术及应用	(187)
6.2.2 神经网络型挖掘技术	(192)
6.2.3 遗传算法型挖掘技术	(197)
6.2.4 粗糙集挖掘技术	(201)
6.2.5 决策树型挖掘技术	(203)
6.3 知识发现工具与应用	(207)
6.3.1 知识挖掘工具的系统结构	(207)
6.3.2 知识挖掘工具运用中的问题	(208)
6.3.3 知识挖掘的价值	(211)
6.3.4 现代数据挖掘工具简介	(211)
6.4 数据挖掘技术的发展	(212)
6.4.1 文本挖掘	(212)
6.4.2 可视化数据挖掘技术	(215)
6.4.3 空间数据挖掘	(216)
6.4.4 分布式数据挖掘	(217)
本章小结	(219)
自我测试题	(220)
第 7 章 电子商务中的商务智能	(223)
7.1 电子商务的数据仓库开发	(224)
7.1.1 电子商务的数据环境分析	(224)
7.1.2 电子商务数据仓库项目的规划	(225)
7.1.3 电子商务数据仓库需求分析	(226)
7.1.4 电子商务数据仓库逻辑模型设计	(226)
7.1.5 电子商务数据仓库的物理模型设计	(227)
7.1.6 电子商务数据仓库的实现与部署	(227)
7.2 电子商务数据仓库的逻辑模型	(228)
7.2.1 营销数据仓库模型演变	(228)
7.2.2 客户维——用户维	(229)
7.2.3 时间维——财政时间维和用户时间维	(230)
7.2.4 地理位置维——物理地理、Web 地理和站点地理维	(231)
7.2.5 商品维——内容维和活动维	(232)
7.2.6 促销维——内部促销维和外部促销维	(233)
7.2.7 电子商务数据仓库模型的事实表设计	(234)
7.3 电子商务数据仓库的 ETL 设计	(234)

7.3.1	电子商务数据仓库的 ETL 环境	(235)
7.3.2	电子商务数据仓库的 ETL 体系结构	(236)
7.3.3	建立 ETL 系统	(237)
7.4	电子商务数据挖掘	(240)
7.4.1	Web 数据挖掘技术	(240)
7.4.2	Web 数据挖掘在电子商务中的应用	(244)
7.4.3	数据挖掘在网站设计、管理中的应用	(245)
7.4.4	数据挖掘在网络广告中的应用	(248)
	本章小结	(250)
	自我测试题	(250)
第 8 章	商务智能应用与管理	(253)
8.1	商务智能用户	(254)
8.1.1	商务智能的用户——信息使用者与知识挖掘者	(254)
8.1.2	信息使用者的商务智能使用方式	(254)
8.1.3	知识挖掘者的商务智能使用方式	(255)
8.2	管理决策中的商务智能应用	(256)
8.2.1	商务智能与供应链管理	(256)
8.2.2	商务智能与客户关系管理	(257)
8.2.3	商务智能和企业资源规划	(259)
8.2.4	商务智能的行业应用	(260)
8.3	商务智能的运行管理	(264)
8.3.1	商务智能中的人力资源管理	(264)
8.3.2	数据加载的管理问题	(265)
8.3.3	故障恢复管理	(266)
8.3.4	访问控制与安全管理	(266)
8.3.5	数据增长的管理	(267)
8.4	商务智能的元数据管理	(268)
8.4.1	元数据的存储、管理和维护	(268)
8.4.2	元数据的用户与使用方法	(270)
8.4.3	元数据的管理模型	(271)
8.5	商务智能中的法律问题	(274)
8.5.1	数据的隐私权问题	(274)
8.5.2	数据隐私权的处理	(275)
8.6	商务智能的成本与效益分析	(276)
8.6.1	商务智能投资回报的定量分析	(277)
8.6.2	商务智能投资回报的定性分析	(278)
	本章小结	(279)
	自我测试题	(279)
第 9 章	商务智能开发实例	(283)
9.1	基于商务智能的超市营销系统的规划与分析	(284)

9.1.1	基于商务智能的超市营销系统的需求分析	(284)
9.1.2	基于商务智能的超市营销系统数据仓库 E-R 模型构造	(285)
9.1.3	基于商务智能的超市营销系统数据仓库事实表模型	(286)
9.1.4	基于商务智能的超市营销系统数据仓库维表模型的设计	(288)
9.1.5	基于商务智能的超市营销系统数据仓库模型的关键字设计	(292)
9.1.6	基于商务智能的超市营销系统数据仓库的元数据设计	(294)
9.2	商务智能开发工具简介	(297)
9.2.1	商务智能开发工具	(297)
9.2.2	SQL Server 商务智能开发应用工具	(301)
9.3	SQL Server 的数据仓库创建	(303)
9.3.1	创建数据库	(303)
9.3.2	创建表	(304)
9.4	SQL Server 数据仓库事实表与多维数据集的建立	(305)
9.4.1	SQL Server BI Development Studio 数据源的确定	(306)
9.4.2	SQL Server 数据仓库数据维的创建	(309)
9.4.3	SQL Server 的多维数据集的创建	(313)
	本章小结	(319)
	自我测试题	(319)
第 10 章	商务智能应用实例	(323)
10.1	数据仓库的数据加载与钻取	(324)
10.1.1	数据仓库的数据加载	(324)
10.1.2	超市营销系统数据仓库的数据加载	(326)
10.1.3	多维数据集的更新	(332)
10.1.4	数据仓库的钻取访问	(334)
10.1.5	数据仓库的多维表达式 MDX 的应用	(337)
10.2	商务智能中的数据挖掘模型设计	(339)
10.2.1	数据挖掘对象的分析	(339)
10.2.2	数据挖掘模型与相关数据的准备	(340)
10.2.3	数据挖掘模型的应用	(342)
10.3	SQL Server 中的数据挖掘工具	(343)
10.3.1	决策类数据挖掘工具的应用	(344)
10.3.2	聚类分析的数据挖掘工具应用	(349)
10.4	商务智能客户端界面的设计	(354)
10.4.1	确定客户端界面展现的内容	(354)
10.4.2	选择客户端界面展现工具	(355)
10.4.3	应用 Excel 实现客户端界面的展现	(356)
	本章小结	(361)
	自我测试题	(362)
参考文献	(365)

第1章

商务智能基本概念

信息技术的不断推广应用,将企业带入了一个信息爆炸的时代。每时每刻都有潮水般的信息出现在管理者的面前,等待管理者去处理、去使用。与此同时,企业的管理者在管理中面对来自不同部门的、相互矛盾的信息无法对所要解决的决策问题提出正确的解决方案。为此,需要一种新的信息处理技术能够使决策者们获取及时准确的信息,以理解商务活动并做出智能化的、更有效的决策,即能从海量的数据中提取有用的信息并转化为商务知识,从而告别“拍脑袋”决策方式。

通过本章的学习,可以了解:

- 商务智能的发展及体系结构;
- 数据仓库的总体结构框架;
- 数据仓库的功能结构;
- 数据仓库的环境支持结构;
- 数据挖掘的基本原理;
- 数据挖掘的应用范围和应用过程。