

商务智能与 医院科学管理



编著：曹琴英 张才明 石仲仁 张廷济 周莲茹

天津出版传媒集团

 天津科学技术出版社



曹琴英，中共党员，主任医师。石家庄市第四医院（妇产医院）院长，党委副书记。中华医学会医学遗传学专业委员会临床遗传学组委员、中国女医师协会常务理事、中国优生优育协会理事、河北省医学会医学遗传学专业委员会副主任委员、河北省医学会妇产科专业委员会常务委员、河北省医院协会常务理事等职务。编写医学论著3部、发表国家级核心期刊论文23篇。获得河北省科技进步三等奖两项、河北省中医药学会科学技术一等奖一项、河北省科学技术成果证书数项。

商务智能与 医院科学管理

编著：曹琴英
张才明
石仲仁
张廷济
周莲茹

天津出版传媒集团



天津科学技术出版社

图书在版编目(CIP)数据

商务智能与医院科学管理 / 曹琴英编著等. -- 天津:
天津科学技术出版社, 2014.2

ISBN 978-7-5308-8760-8

I. ①商… II. ①曹… III. ①电子商务—应用—医院—
—管理 IV. ①R197.324

中国版本图书馆CIP数据核字(2014)第020269号

责任编辑:丁红敬

策划编辑:王 冬

天津出版传媒集团



天津科学技术出版社

出版人:蔡 颢

天津市西康路 35 号 邮编 300051

电话(022)23332377

网址:www.tjkjcb.com.cn

新华书店经销

天津市方正汇智彩色印刷技术有限公司

开本 850×1168 1/32 印张 15.75 字数 342 000

2014 年 2 月第 1 版第 1 次印刷

定价:36.00 元

编委会成员

主 编：曹琴英
张才明
石仲仁
张廷济
周莲茹

副主编：任怀志
刘兰平
吕 辉
黎安明
王长岱
王彦博
孙国志
赵媛媛
张 雷
李 轶

内 容 提 要

本书介绍商务智能概念、商务智能和医院战略管理理论、医院科学管理模式，以及支撑和实现医院科学管理模式的信息化战略和商务智能应用，对具体的商务智能在医院科学管理中的各个领域的应用也做了相应介绍。最后采用实证案例和解决方案，给广大医院实现科学管理提供实际的参考和路径选择。

目 录

第一章 数据挖掘与商务智能概述	(1)
第一节 数据挖掘概述	(1)
第二节 商务智能概述	(19)
第三节 数据挖掘与商务智能	(36)
第二章 医院信息化发展现状及发展方向	(53)
第一节 医院信息化与医院信息系统	(53)
第二节 国际医院行业信息化	(58)
第三节 国内医院行业信息化	(63)
第三章 医院科学管理	(71)
第一节 医院科学管理理论	(71)
第二节 医院科学管理模式	(89)
第三节 医院科学发展理论	(105)
第四节 医院科学战略理论	(117)
第四章 商务智能在医院科学管理中的应用	(133)
第一节 商务智能在政府卫生行业管理部门中的应用 ...	(133)
第二节 商务智能在区域卫生医疗平台中的应用	(137)
第三节 商务智能在医院科学管理与决策中的应用	(140)
第四节 商务智能在医院财务科学管理中的应用	(148)
第五节 商务智能在医院成本科学管理中的应用	(156)
第六节 商务智能在医院绩效科学管理中的应用	(180)
第七节 商务智能在医院住院科学管理中的应用	(205)
第八节 商务智能在医院手术科学管理中的应用	(206)
第九节 商务智能在医院病案科学管理中的应用	(207)

第十节	商务智能在医院药品科学管理中的应用	·····	(209)
第十一节	商务智能在医院患者科学管理中的应用	·····	(213)
第十二节	商务智能在医院后勤科学管理中的应用	·····	(213)
第十三节	商务智能在医院质量科学管理中的应用	·····	(214)
第五章	商务智能在医院科学管理中的实证	·····	(218)
第一节	实证方案: Diver医院商务智能解决方案	·····	(218)
第二节	实证案例: 河北省儿童医院	·····	(236)

第一章 数据挖掘与商务智能概述

第一节 数据挖掘概述

一、数据挖掘的基本概念

随着数据库、网络等技术的迅速发展,人们积累的数据越来越多,我们已经被淹没在数据和信息的汪洋大海中。人们需要有新的、更有效的方法对各种大量的数据进行分析、提取以挖掘其潜能,数据挖掘正是在这样的应用需求环境下产生并迅速发展起来的,它的出现为把海量数据智能转化为有用的信息和知识提供了新的思路 and 手段。

数据挖掘 (Data Mining),指的是从大型数据库或数据仓库中提取人们感兴趣的知识,这些知识是隐含的、事先未知的潜在有用信息。数据挖掘是一个高级的处理过程,它从数据集中识别出以模式来表示的知识。高级的处理过程是指一个多个步骤的处理过程,多个步骤之间相互影响、反复调整,形成一种螺旋式上升过程。

数据挖掘也是一个多学科交叉研究领域,它融合了数据库技术、人工智能、机器学习、统计学、知识工程、面向对象方法、信息检索、高性能计算以及数据可视化等最新技术的研究成果。数据挖掘就是从大量存储的数据中,利用模式识别、统计和数学的技术,筛选发现新的有意义的关系、模式和趋势的方法。就是从大量的、不完全的、有噪声的、模糊的、随机的数据中,提取隐含在其中的、人们事先不知道的、但又是潜在有用的信息和知识的过程。它主要依靠人工智能、机器学习和统计学技术,对数据进行归纳推理,从中挖掘出潜在的模式,预测未来趋势,为决策提供支持。

起初,各种商业数据仅仅存储在数据库中,然后逐渐到对数据库

中的商业数据进行查询,进而发展到对数据的即时遍历。数据挖掘使数据库技术的应用进入到更高级的阶段,它不仅能对过去的数据进行查询,并且能够找出数据之间的内在关系,从而加大信息应用的深度。随着海量数据搜集、强大的多处理器计算机和数据挖掘算法这三种基础技术的发展成熟,数据挖掘技术在商业应用中开始得到广泛的重视。

数据挖掘的定义很多,表达方式各不相同。从技术角度看,数据挖掘是从大量的、不完全的、有噪声的、模糊的、随机的实际数据中,提取隐含在其中的、人们不知道的、但又是潜在有用的信息和知识的过程;从商业应用角度看,数据挖掘是一种崭新的商业信息处理技术。其主要特点是对商业数据库中的大量业务数据进行抽取、转化、分析和模式化处理,从中提取辅助商业决策的关键知识,即从一个数据库中自动发现相关商业模式;数据挖掘是从存放在数据库、数据仓库或其他信息库中的大量数据中挖掘有趣知识的过程。

与数据挖掘相近的同义词有数据融合、数据分析和决策支持等。数据挖掘技术是人们长期对数据库技术进行研究和开发的结果。起初各种商业数据是存储在计算机的数据库中的,然后发展到可对数据库进行查询和访问,进而发展到对数据库的即时遍历。数据挖掘使数据库技术进入了一个更高级的阶段,它不仅能对过去的数据进行查询和遍历,并且能够找出过去数据之间的潜在联系,从而促进信息的传递。现在数据挖掘技术在商业应用中已经得到成熟使用,因为对这种技术进行支持的三种基础技术已经发展成熟,它们是:海量数据搜集、强大的多处理器计算机、数据挖掘算法。

数据挖掘与传统的数据分析(如查询、报表、联机应用分析)的本质区别是数据挖掘是在没有明确假设的前提下去挖掘信息、发现知识。数据挖掘所得到的信息应具有“先前未知”“有效”和“可实用”三个特征。其中,“先前未知”的信息是指该信息是预先未曾预料到的,即数据挖掘是要发现那些不能靠直觉发现的信息或知识,甚至是违背直觉的信息或知识,挖掘出的信息越是出乎意料,就可能越

有价值。在商业应用中最典型的例子就是一家连锁店通过数据挖掘发现了小孩尿布和啤酒之间有着惊人的联系。数据仓库技术的发展与数据挖掘有着密切的关系。数据仓库的发展是促进数据挖掘越来越热的原因之一。但是,数据仓库并不是数据挖掘,在人工智能领域,习惯上又称为数据库中的知识发现,也有人把数据挖掘视为数据库中知识发现过程的一个基本步骤。知识发现过程由以下三个阶段组成:(1)数据准备,(2)数据挖掘,(3)结果表达和解释。数据挖掘可以与用户或知识库交互。数据挖掘是通过分析每个数据,从大量数据中寻找其规律的技术,主要有数据准备、规律寻找和规律表示 3个步骤。数据准备是从相关的数据源中选取所需的数据并整合成用于数据挖掘的数据集;规律寻找是用某种方法将数据集所含的规律找出来;规律表示是尽可能以用户可理解的方式(如可视化)将找出的规律表示出来。数据挖掘的任务有关联分析、聚类分析、分类分析、异常分析、特异群组分析和演变分析,等等。

并非所有的信息发现任务都被视为数据挖掘。例如,使用数据库管理系统查找个别的记录,或通过因特网的搜索引擎查找特定的 Web 页面,则是信息检索领域的任务。虽然这些任务是重要的,可能涉及使用复杂的算法和数据结构,但是它们主要依赖传统的计算机科学技术和数据的明显特征来创建索引结构,从而有效地组织和检索信息。尽管如此,数据挖掘技术也已用来增强信息检索系统的能力。(见图 1)

1、数据挖掘技术

数据挖掘,就是从大量的、不完全的、有噪声的、模糊的、随机的数据库中,提取隐含在其中的、人们事先不知道的、但又是潜在有用的信息和知识的过程。

2、数据挖掘过程

数据挖掘过程包括很多处理阶段,其一般流程主要包括三个阶段:数据准备、数据挖掘、结果解释和评价。

1)数据准备:数据准备又可以分为两个子阶段:数据选取、数据预处理。数据选取的目的是确定发现任务的操作对象,即目标数据。

它是根据用户的需要从原始数据库中抽取的一组数据。数据预处理一般包括消除噪声、推导计算缺值数据、消除重复记录、完成数据类型转换 (如把连续型数据转换为离散型数据, 以便于符号归纳; 或是把离散性数据转换为连续型数据, 以便于神经网络计算) 以及对数据进行降维处理 (即从初始特征中找出真正有用的特征以减少数据挖掘要考虑的变量个数)。

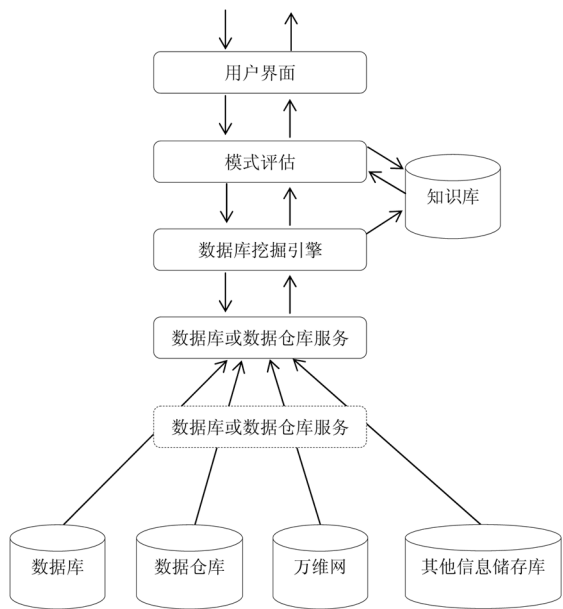


图 1 典型数据挖掘系统的结构

2)数据挖掘:数据挖掘阶段首先要确定数据挖掘的目标和挖掘的知识类型;然后根据挖掘的知识类型选择合适的挖掘算法;最后实施数据挖掘操作,运用选定的挖掘算法从数据库中抽取所需的知识。

3)结果的解释和评价:数据挖掘阶段发现的知识经过评估,可能存在冗余或无关的知识,这时需要将其剔除;也有可能知识不满足用户的要求,需要重复上述挖掘过程重新进行挖掘。另外,由于数据

挖掘是最终要面临用户的,因此,还需要对所挖掘的知识进行解释,以一种用户易于理解的方式(如可视化方式)供用户所用。

可以看出,以上整个数据挖掘过程是不断地循环和反复的,因而可以对所挖掘出来的知识不断求精和深化,最终达到用户所满意的结果。

二、数据挖掘的基本步骤

数据挖掘的步骤会随不同领域的应用而有所变化,每一种数据挖掘技术也会有各自的特性和使用步骤,针对不同问题和需求所制定的数据挖掘过程也会存在差异。此外,数据的完整程度、专业人员支持的程度等都会对建立数据挖掘过程有所影响。这些因素造成了数据挖掘在各不同领域中的运用、规划,以及流程的差异性,即使同一产业,也会因为分析技术和专业知识的涉入程度不同而不同,因此对于数据挖掘过程的系统化、标准化就显得格外重要。如此一来,不仅可以较容易地跨领域应用,也可以结合不同的专业知识,发挥数据挖掘的真正精神。

数据挖掘完整的步骤如下:

理解数据和数据的来源 (understanding)。

获取相关知识与技术 (acquisition)。

整合与检查数据 (integration and checking)。

去除错误或不一致的数据 (data cleaning)。

建立模型和假设 (model and hypothesis development)。

实际数据挖掘工作 (data mining)。

测试和验证挖掘结果 (testing and verification)。

解释和应用 (interpretation and use)。

由上述步骤可看出,数据挖掘牵涉了大量的准备工作与规划工作,事实上许多专家都认为整套数据挖掘的过程中,有 80% 的时间和精力是花费在数据预处理阶段,其中包括数据的净化、数据格式转换、变量整合,以及数据表的链接。可见,在进行数据挖掘技术的分析之前,还有许多准备工作要完成。(见图 2)

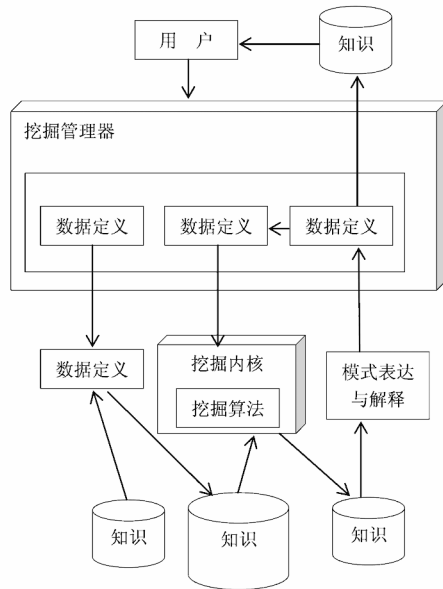


图 2 一个数据挖掘系统原型

三、数据挖掘的一般流程

数据挖掘的过程一般由六个阶段组成,这些阶段之间的顺序并不是线性的,为了取得好的数据挖掘结果常常有必要反复重复这些步骤,究竟下一步要执行哪个阶段或者哪一个特定的任务,取决于每一个阶段的结果。

1、确定业务对象

在开始数据挖掘之前最重要的一步就是定义业务对象和了解相关资料。挖掘的最后结果是不可预测的,但要探索的问题应是有预见的。如果缺少对待解决问题的明确定义和相关背景知识,就不能为挖掘准备资料,也很难正确地解释得到的结果。因此要充分发挥数据挖掘的价值,必须要对目标有清晰明确的定义,有对数据挖掘项目得到的结果进行衡量的标准和对整个项目预算合理性的解释。

2、数据准备

数据准备包括了对数据的选择、预处理和清洗等工作。确定业

务对象后,需要搜索所有与业务对象有关的内部和外部数据信息,并从中选择出适用于数据挖掘用的数据。然后研究所选数据的质量,保证其适合于挖掘应用,为进一步的分析作准备。在选择数据后,要对数据进行预处理和清洗,解决数据的冗余、缺值,据定义不一致等问题。此外,为了得到更好的结果,数据准备和数据挖掘工作可能需要不断的反复进行。因此,数据准备工作大概要花去整个数据挖掘项目的 50% – 90% 的时间和精力。

3、数据的转换

将准备好的数据转换成一个真正适合挖掘算法的分析模型是数据挖掘成功的关键。模型的建立从数据分析开始,首先选择变量,然后从原始数据中构建新的预示值,接着从数据中选取一个子集或者样本来建立模型,最后转换变量,使之和建立模型的算法一致。建立模型是一个反复的过程,要仔细考察不同的模型以判断哪个模型对定义的问题最有用。

4、数据挖掘

对所得到的经过转换的数据进行挖掘。除了选择合适的挖掘算法外,其余工作一般由挖掘工具自动完成。

5、结果解释与评估

当得出数据挖掘结果后,必须评价它的结果、解释它的价值。使用的分析方法一般应根据数据挖掘操作而定,通常会用到可视化技术。

6、知识的发现和应用

数据挖掘的结果经验证之后,可以提供给分析人员做参考,由他们通过察看和分析结果之后提出行动方案建议。还可以将分析所得到的知识集成到业务信息系统的组织结构中去,使数据挖掘得到的知识发现可以在决策分析中得到应用。

四、常用的数据挖掘技术

常用的数据挖掘技术可以分成统计分析类、知识发现类和其他类型的数据挖掘技术三大类。

1、统计分析类

统计分析(或称数据分析)技术中使用的数据挖掘模型有线性分析和非线性分析、回归分析、逻辑回归分析、单变量分析、多变量分析、时间序列分析、最近邻算法和聚类分析等技术。利用这些技术可以检查那些异常的数据,然后利用各种统计模型和数学模型解释这些数据,解释隐藏在这些数据背后的市场规律和商业机会。例如,可以使用统计分析工具寻找最佳商业机会,增加市场份额及利润;利用质量管理程序,提高产品或服务的质量,使客户更加满意;通过对产品制造的调整或企业业务重组,增加利润。在所有的数据挖掘技术中,统计型数据挖掘技术是最成熟的一种,已经在数据挖掘中得到广泛的应用。

2、知识发现类

知识发现类数据挖掘技术是与统计类数据挖掘技术完全不同的一种挖掘技术。它可以从数据仓库的大量数据中筛选信息,寻找市场中可能出现的新的运营模式,发掘出人们所不知道的事实。知识发现类数据挖掘技术包含人工神经网络、决策树、遗传算法、粗糙集、规则发现和关联顺序等。

3、其他数据挖掘技术

其他数据挖掘技术中包含文本数据挖掘、Web数据挖掘、分类系统、可视化系统、空间数据挖掘和分布式数据挖掘等。

(1)文本数据挖掘是为了进行非结构化信息的挖掘;

(2)Web数据挖掘是针对 Internet技术所带来的大批量的网络信息的挖掘;

(3)可视化系统使数据挖掘以图形或图像的方式在屏幕上显示出来,且能交互处理,能直观地发现隐含有用的知识;

(4)空间数据挖掘是基于地理信息系统的数据挖掘技术,可用于地图、预处理后的遥感数据、医学图像数据和芯片设计空间数据库中非显式的知识、空间关系和其他有意义的模式的提取;

(5)分布式数据挖掘更有利于对分布式数据库数据资源的利用。

五、数据挖掘的主要方法

在数据挖掘的处理过程中,数据挖掘方法是最为关键的,而目前数据挖掘的方法主要有以下几类:

(1)关联规则方法:关联规则挖掘就是发现存在于大量数据集中的关联性或相关性,关联规则本身是一种很重要的知识,是数据挖掘研究的主要内容,它能表示数据之间的相互关系,对统计和决策工作有重大意义。

(2)分类和聚类方法:分类就是假定数据库中的每个对象(在关系数据库中对象是元组)属于一个预先给定的类,从而将数据库中的数据分配到给定的类中。而聚类是将地实体对象集合依照某种相似性度量原则划分为若干个类似实体对象组成的多个类或簇的过程。分类和聚类都是对目标进行划分,划分的标准是类内差别最小,而类间差别最大。分类和聚类的区别在于,分类事先知道类别数和各类的典型特征,而聚类则事先不知道。

(3)数据统计方法:使用这些方法一般首先建立一个数据模型或统计模型,然后根据这种模型提取有关的知识。例如,可由训练数据建立一个 Bayesian 网,然后,根据该网的一些参数及联系权重提取出相关的知识。

(4)机器学习方法:大多数机器学习方法使用人类的认识模型模仿人类的学习方法从数据中提取知识,由于机器学习经过多年的研究,已取得了一些较满意的成果,因此,在数据挖掘中可以利用目前比较成熟的机器学习方法。

(5)多层次数据汇总归纳:数据库中的数据和对象经常包含原始概念层上的详细信息,将一个数据集合归纳成更高概念层次信息的数据挖掘技术被称为数据汇总。概念汇总是将数据库中的相关数据由低概念层抽象到高概念层,主要有数据立方体和面向对象的归纳两种方法。

(6)神经网络方法:神经网络本身具有良好的鲁棒性、自组织自适应性、并行处理、分布存储和高度容错等特性,非常适合解决数据挖掘的问题,近年来越来越受到人们的关注。典型的神经网络模型