第一图书网, tushu007.com

<<Clementine数据挖掘方法及应用>>

图书基本信息

书名：<<Clementine数据挖掘方法及应用>>

13位ISBN编号：9787121117787

10位ISBN编号：7121117789

出版时间：2010-9

出版时间：电子工业

作者：薛薇//陈欢歌

页数：303

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

第一图书网, tushu007.com

前言

数据挖掘是当前数据分析领域中最活跃最前沿的地带。
任何事物都有定性和定量两个方面，定量则产生数据。
从数据分析入手是我们认识事物本质的基本手段。
任何事物都是互相关联着的，从数据分析入手是我们把握事物之间联系的基本方法。
任何事物都在永恒地变化发展着，从数据分析入手是我们探索事物发展规律的基本思路。
所以我们进行数据分析，既是一种世界观，也是一种方法论。
我们在研究着丰富多彩的客观世界的同时，也体现着分析者主观的智慧和自身的价值。
随着中国社会经济的蓬勃发展，在错综复杂的宏观、中观和微观的共同作用下，战略决策和战术选择都显得敏感而关键，越来越多的人们加入到数据分析的行列中来。
这是一个非常富有挑战性的工作，不但有意思而且有意义。
IBM公司于2009年1月公布了其“智慧地球”战略。
该战略的主要思想是，将传感设备或智能仪表嵌入到建筑、电力、交通、管道等各种物体中，进行数据自动采集，之后基于互联网形成物物相联的物联网，然后通过超级计算机和云计算将数据整合，进行智能化分析和建模，从而实现社会与物理世界的融合。
这是一个未来理想化的信息世界图景。
在这个智慧系统中，其核心是数据处理。
为此，IBM公司于2009年7月斥资12亿美元收购了著名的SPSS统计分析软件公司，将其应用广泛的SPSS统计分析软件和Clementine数据挖掘软件纳入麾下。
同时对软件产品进行了整合，将Clementine更新命名为PASW（Predictive Analytics Software）Modeler，并快速推向市场。
目前，SPSS Clementine软件已经连续若干年蝉联数据挖掘应用的王者，而业界对于PASW Modeler的认知则刚刚开始。
所以本书继续沿用为广大读者所熟悉的Clementine这个名字。
Clementine软件不但将计算机科学中许多机器学习的优秀算法带入到数据分析中来，同时也综合了一些行之有效的数据挖掘方法，成为内容最为全面、功能最为强大的数据挖掘产品。
Clementine软件充分利用计算机系统的运算处理能力和图形展现能力，将方法、应用与工具有机地融合为一体，是解决数据挖掘问题的最理想工具。
Clementine软件继续保持了SPSS产品的一贯风格：界面友好且容易使用。
复杂的数学算法和冗余的输出结果被软件隐藏在程序系统内部。

第一图书网, tushu007.com

<<Clementine数据挖掘方法及应用>>

内容概要

数据挖掘是当前数据分析领域中最活跃最前沿的地带。
本书以数据挖掘的实践过程为主线，通过生动的应用案例，从数据挖掘实施角度，系统介绍了经典的数据挖掘方法和利用Clementine实现数据挖掘的全部过程，讲解方法从易到难，说明问题从浅至深。
本书力求以最通俗的方式阐述数据挖掘方法的核心思想与基本原理，同时配合Clementine软件操作的说明，希望读者能够直观了解方法本质，尽快掌握Clementine软件使用，并应用到数据挖掘实践中。
为方便读者学习，书中所有数据和案例与所附光盘内容一致。
本书适合于从事数据分析各应用领域的读者，尤其适合于商业管理、财政经济、金融保险、社会研究、人文教育等行业的相关人员。
同时，也能够作为高等院校计算机类、财经类、管理类专业本科生和研究生的数据挖掘教材。

第一图书网, tushu007.com

<<Clementine数据挖掘方法及应用>>

书籍目录

第一图书网, tushu007.com

第一图书网, tushu007.com

章节摘录

插图：数据挖掘，作为20世纪90年代中后期兴起的，具有鲜明跨学科色彩的应用和研究领域，因其注重减少数据分析方法对数据的限制性和约束性，注重与计算机技术结合以实现数据的可管理性以及分析的易操作性，已成为数据分析应用实践的新生代。
同时，随着数据挖掘方法的不断成熟及其应用的日益普及化，数据挖掘软件的研发也取得了令人可喜的成果。
目前，以Clementine为代表的数据挖掘软件，因其有效地将束之高阁的数据挖掘理论成果解放到数据分析实践中，已普遍应用于商业、社会、经济、教育、金融、医学等领域，并成为数据分析的主流工具，得到数据分析相关领域的极大关注。
1.1 数据挖掘的产生背景数据挖掘的产生和兴起是在计算机数据库技术蓬勃发展，人工智能技术应用领域不断拓展，统计分析方法不断丰富过程中，为有效迎合数据分析的实际需求而逐步形成和发展起来的一门具有鲜明跨学科色彩的应用研究领域。
1.1.1 海量数据的分析需求催生数据挖掘20世纪80年代以来，随着计算机数据库技术和产品的日益成熟以及计算机应用的普及深化，各行业部门的数据采集能力得到了前所未有的提高，组织通过各自内部的业务处理系统、管理信息系统以及外部网络系统，获得并积累了浩如烟海的数据。
以商业领域为例，美国著名的连锁超市Wal-Mart的数据库中已积累了TB级以上的顾客购买行为数据和其他销售数据。
随着互联网和电子商务的普及，各类网上书店、网上银行、网上营业厅和网上商城等积累的Web点击流数据，存储容量也多高达GB级。
另外，国家政府部门所积累的数据量也令人瞠目。
例如，一次全国经济普查或人口普查所采集和处理数据量均在千万级以上。
同时，各经济行业的企业内部也拥有大量的业务数据、财务数据和人事数据。
在严酷的市场竞争压力下，企业为更客观地把握自身和市场状况，提升内部管理和决策水平，管理者们面对如此丰富的海量数据，分析需求越来越强烈。

第一图书网, tushu007.com

编辑推荐

《Clementine数据挖掘方法及应用》由电子工业出版社出版。

第一图书网, tushu007.com

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