

<<数据挖掘>>

图书基本信息

<<数据挖掘>>

内容概要

数据挖掘是信息领域发展最快的技术，很多不同领域的专家，比如统计学家、数据库专家等，都从中获得了发展的空间，使得数据挖掘日益成为企业界讨论的热门话题。

随着信息技术的发展，人们采集数据的手段日益丰富与高明，由此积累的数据日益膨胀，数据量达到GB甚至TB级，而且高维数据也日益成为主流。

这些海量数据及其高维特征使得传统的数据分析手段相形见绌。

计算机性能的日益更新，使得人们能够期望计算机帮助我们分析与理解数据，帮助我们以丰富的数据为基础做出正确决策。

<<数据挖掘>>

作者简介

迈克尔·贝里和戈登·利诺夫是数据挖掘公司的创始人，这是美国一家很受欢迎的数据挖掘顾问公司。在从事数据挖掘项目之余，他们在世界各地讲课、演讲，所到之处均大受欢迎。

公司站点：www.data-miners.com，该站点内容包括：最新的数据挖掘产品和服务提供商的信息

<<数据挖掘>>

书籍目录

第1部分 本书焦点 第1章 数据挖掘概述 第2章 为什么要精通数据挖掘这门艺术 第3章 数据挖掘方法论：互动循环系统 第4章 客户和他们的生命周期第2部分 数据挖掘的三大支柱 第5章 数据挖掘技术与算法 第6章 无所不在的数据 第7章 建立有效的预测模型 第8章 实施控制：建立数据挖掘环境第3部分 案例研究 第9章 数据挖掘在目录直销业中的应用 第10章 数据挖掘在在线银行业中的应用 第11章 数据挖掘在无线通信业中的应用 第12章 数据挖掘在电信业中的应用 第13章 谁正在买什么？
第14章 不浪费、不短缺：改善生产流程 第15章 社会议题：数据挖掘与隐私权 索引

章节摘录

第1部分 本书焦点 第1章 数据挖掘概述 要准确地回顾并概括过去一个世纪的巨变是很难的事情。

尽管我们可以直接感受到各个领域 中的变化，但我们很难发现哪个领域的变化能超越信息数量的增加。

信息爆炸为各行各业提供了新的机遇，同时也引致了一些新的问题，从制造业到医药行业再到市场营销无不如此。

若要评价过去若干年世界信息储量的增加，就有必要回顾20世纪中具有历史意义的事件。

1900年，世界人口是16亿（除信息之外另一个可称为“爆炸”的领域）。

而100年后，世界人口达到了60亿。

人口爆炸导致在20世纪结束时人口达到该世纪初的3.75倍。

1906年，斯坦利孪生兄弟——弗朗西斯和弗里兰（Stanley Francis和Stanley Freelan）用斯坦利蒸汽机车创造了每小时122英里的地速世界记录。

在当时，船在水中每小时行驶15英里已经是相当快的了；而飞机能够上天也才仅有3年的历史。

因而，当时地速的这一记录没有挑战者。

而在63年以后，阿波罗号宇宙飞船飞往月球时的速度达到了每小时25000英里，是蒸汽机车创造的记录的205倍。

登月飞行是另一个新的里程碑。

1900年，人们所能抵达的最远的旅程大约为25000英里，也就是环绕地球一圈的距离。

到达月亮的往返路程大约是环球旅行的19倍。

以上这些数字给人的印象是深刻的，但所有这些都无法与公司数据的增长相比。

在20世纪初，或者即使到了20世纪中叶，没有一个公司的账目、订货记录和文件柜的数据总和能超过几十个百万字节（megabytes）。

今天，最大的公司数据库的容量是用万亿字节（terabytes）来计量的。

对于这些公司来说，数据的增长是以10万倍来计量的。

这些比较主要是用文字和数字的形式进行（本书都采用相似的比较）。

影像和录音也可储存大量数据和信息，同样是我们所关注的，但本书所介绍的数据挖掘技术还不能适用于视频和音频数据的挖掘工作中。

<<数据挖掘>>

媒体关注与评论

“在21世纪，公司的生存取决于海量商务数据是否被有效利用。
贝里和利诺夫带领读者穿越了一条（数据挖掘）最佳实践的启迪之路。

”——吉姆·古德奈特，SAS研究所主席和联合发起人 “数据挖掘是理解网上客户行为，并帮助互联网企业创建个性化客户服务的重要工具，是电子商务时代的重要工具。
这是一本伟大的书，它是我专业工作中四、五本必备手册之一。

”——拉尔夫·金博尔，《数据仓库生命周期工具箱》作者 “本书解决了二个重要问题：从实际操作层面阐述数据挖掘并将它与商业世界联系起来，太多数据挖掘者忘记了他们最终是为商业服务的，但本书的作者没有犯这个低级错误。

如果你对数据挖掘有兴趣，本书是必读书。

”——比尔·因蒙，《创建数据仓库》作者

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>