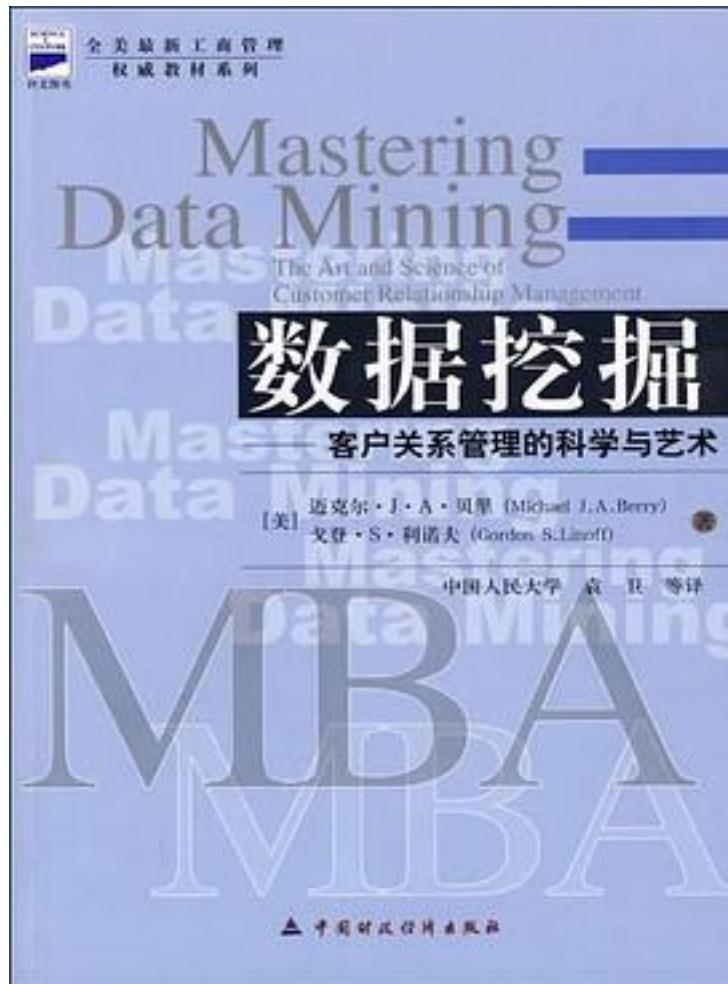


数据挖掘



[数据挖掘 下载链接1](#)

著者: (美) Jiawei Han

出版者:机械工业出版社

出版时间:2012-3

装帧:

isbn:9787111374312

数据挖掘领域最具里程碑意义的经典著作

完整全面阐述该领域的重要知识和技术创新

【编辑推荐】

我们生活在数据洪流的时代。本书向我们展示了如何从这样海量的数据中找到有用知识的方法和技术。最新的第3版显著扩充了数据预处理、挖掘频繁模式、分类和聚类这几个核心章节的内容；还全面讲述了OLAP和离群点检测，并研讨了挖掘网络、复杂数据类型以及重要应用领域。本书将是一本适用于数据分析、数据挖掘和知识发现课程的优秀教材。

—— Gregory Piatetsky-Shapiro, KDnuggets的总裁

Jiawei、Micheline和Jian的教材全景式地讨论了数据挖掘的所有相关方法，从聚类和分类的经典主题，到数据库方法（关联规则、数据立方体），到更新和更高级的主题（SVD/PCA、小波、支持向量机），等等。总的说来，这是一本既讲述经典数据挖掘方法又涵盖大量当代数据挖掘技术的优秀著作，既是教学相长的优秀教材，又对专业人员具有很高的参考价值。

—— 摘自卡内基-梅隆大学Christos Faloutsos教授为本书所作序言

【内容简介】

当代商业和科学领域大量激增的数据量要求我们采用更加复杂和精细的工具来进行数据分析、处理和挖掘。尽管近年来数据挖掘技术取得的长足进展使得我们广泛收集数据越来越容易，但技术的发展依然难以匹配爆炸性的数据增长以及随之而来的大量数据处理需求，因此我们比以往更加迫切地需要新技术和自动化工具来帮助我们将这些数据转换为有用的信息和知识。

本书前版曾被KDnuggets的读者评选为最受欢迎的数据挖掘专著，是一本可读性极佳的教材。它从数据库角度全面系统地介绍数据挖掘的概念、方法和技术以及技术研究进展，并重点关注近年来该领域重要和最新的课题——数据仓库和数据立方体技术，流数据挖掘，社会化网络挖掘，空间、多媒体和其他复杂数据挖掘。每章都针对关键专题有单独的指导，提供最佳算法，并对怎样将技术运用到实际工作中给出了经过实践检验的实用型规则。如果你希望自己能熟练掌握和运用当今最有力的数据挖掘技术，那这本书正是你需要阅读和学习的宝贵资源。本书是数据挖掘和知识发现领域内的所有教师、研究人员、开发人员和用户都必读的一本书。

本书特点

引入了许多算法和实现示例，全部以易于理解的伪代码编写，适用于实际的大规模数据挖掘项目。

讨论了一些高级主题，例如挖掘面向对象的关系型数据库、空间数据库、多媒体数据库、时间序列数据库、文本数据库、万维网以及其他领域的应用等。

全面而实用地给出用于从海量数据中获取尽可能多信息的概念和技术。

作者介绍：

Jiawei Han(韩家炜) 伊利诺伊大学厄巴纳-尚佩恩分校计算机科学系Abel Bliss教授。由于在数据挖掘和数据库系统领域卓有成效的研究工作，他曾多次获得各种荣誉和奖励，其中包括2004年ACM SIGKDD颁发的最佳创新奖，2005年IEEE Computer Society 颁发的技术成就奖，2009年IEEE颁发的W. Wallace

McDowell奖。他是ACM和IEEE Fellow，同时还是《ACM Transactions on Knowledge Discovery from Data》杂志的主编（2006-2011），以及《IEEE Transactions on Knowledge and Data Engineering》和《Data Mining and Knowledge Discovery》杂志的编委会成员。

Micheline Kamber 拥有加拿大康考迪亚大学计算机科学硕士学位，她是NSERC Scholar，现在加拿大麦吉尔大学、西蒙-弗雷泽大学及瑞士从事研究工作。

Jian Pei (裴健) 目前是加拿大西蒙-弗雷泽大学计算机学院副教授。2002年，他在Jia wei Han教授的指导下获得西蒙-弗雷泽大学博士学位。

目录: Foreword to Second EditionPrefaceAcknowledgmentsAbout the AuthorsChapter 1 IntroductionChapter 2 Getting to Know Your DataChapter 3 Data PreprocessingChapter 4 Data Warehousing and Online Analytical ProcessingChapter 5 Data Cube TechnologyChapter 6 Mining Frequent Patterns, Associations, and Correlations: Basic Concepts and MethodsChapter 7 Advanced Pattern MiningChapter 8 Classification: Basic ConceptsChapter 9 Classification: Advanced MethodsChapter 10 Cluster Analysis: Basic Concepts and MethodsChapter 11 Advanced Cluster AnalysisChapter 12 Outlier DetectionChapter 13 Data Mining Trends and Research FrontiersBibliographyIndex
· · · · · (收起)

[数据挖掘](#) [下载链接1](#)

标签

数据挖掘

计算机

英文版

Data-Mining

DataMining

数据库

硕士教材

数据仓库

评论

封面好棒。老师PPT和课本有些对不上。英文很好懂。

很棒的一本书，就是略难。

数据挖掘入门的佳作，就数据挖掘涉及到的基础知识介绍的很全面，通俗易懂。

这本书的整个架构非常好。

虽然没有完全读完，但是也算是掌握了基本的概念与方法，考完试啦

偏理论，对应用少有提及。

内容较丰富，甚至可以说庞杂，多是概念的介绍，但有浅尝辄止的感觉，不太能找到重点，但作为入门教材其难度还是适当的。

了解数据挖掘这个热门名词的首选书籍，这本书由数据挖掘领域有名的专家编写，能够从大体上把握有关数据的各种概念以及流行技术。

[数据挖掘 下载链接1](#)

书评

浙大的王灿老师的讲课视频：

<http://www.businessanalysis.cn/viewthread.php?tid=13320&extra=&page=1>

韩家炜自己的讲课视频：http://v.youku.com/v_playlist/ct250f1903290o1p0

推荐几个数据分析网站

随着数据的数量级增长，促进了数据分析的火热。但很多数据分析从业人员却感觉没有很好的获取行业资讯、技术交流的平台，今天就推荐几个经常上的觉得不错的网站给大家。大数据时代已经来临，我们一起感受数据的力量吧！数据分析交流群推...

应该说这部书可以把人引进门，但看了之后，总觉得还有些概念模糊之处，比如说数据挖掘的理论来源是什么？如何把这些算法从本质上分类？

我觉得，这方面，《实用数据挖掘》会更好些。另外，如何使用简单的软件，为企业或政府部门实现一个简单可见的数据挖掘呢？这方面，我只读...

首先我承认我的E文不算太好。我最近在读两本E文原版的技术书籍，《分布式系统原理与范型》、《数据挖掘概念与技术》。都是E文书，但是前者作者是老外，后者作者是中国人，两本书的风格明显差异很大。

《数据挖掘概念与技术》的大量专业词汇给我很大的打击，但是大量非专业词汇...

对于刚入门数据挖掘的人来说，这书绝对会让你感觉自己是个折翼的天使。, 因为一开始就各种各样的理论扑面而来，而对于那些经典的算法却只是做一个感性的介绍，并没有那种流程图式的清晰解说。总之就是，不易上手。

但是在这种不善的情况，为什么该书却被国内外...

这本书被翻译的佶屈聱牙，除了给学习数据挖掘的人增添负担，什么积极的作用的没有。不知道有多少人因为这本不通的书而失去对数据挖掘的兴趣。

教授真的是毁人不倦啊，各种官方语言，妈的是要当官吗？

一本引导你入门的书，知识深浅都涵盖，描述广泛但不详实易懂。

前几个chapter屁话较多，但OLAP的概念是有用的。随后的cluster, association的分析解释还是涵盖的很好，但都是点到为止，颇具教科书的味道，其实被来就是一本教科书。剩下的章节就不能看了。6年前就通读此书，...

我了个擦，
连个非限制性定语从句都翻译不了，你翻译毛啊。还不如看原版。你们两个真是叫兽啊。
。本来都不屑去骂，但是连个定于从句都搞不通顺，叫兽你就这水平？你让研究生替你
翻译的话，你研究生的水平也不至于如此奇差吧，还没过四级呢吧。不评很差是看在原
著的面子上。

这本书是准备跟随浙江大学的课程学习而购买的课本，里面的知识比较全面。部分比较深入的知识由于课上没有讲解，因此我也将它跳过了。因为这学期选修了数据挖掘的课，需要一个中文版的课本进行学习，选择这本书还是不错的。
这本书很适合自学，因为是将理论与算法相结合讲解的...

//2017-05-20 13:30

这篇文章我已经欠了至少一年了，周五写记录时，本想写开始认真搞黑客，但突然发现之前的总结少这篇，心里实在过不去，遂补上，顺便梳理一下之前的学习总结，也了却一心愿。

数据挖掘的目标是从数据集中识别出一种或多种模式，并用所发现的模式进行分析或...

大三下时就买了，为了准备一下保研的方向，当时只是粗略的读懂了一点。浙大面试时问了一个K-Means自己都记不太清了。

研一上的<<数据仓库与数据挖掘>>课程也基本使用了这本教材，然而长期不去上课导致自己好多内容学的并不扎实，最后的考试也考的很烂；现在回想，贝叶...

作者是FP-Growth的发明人之一，本身实力不弱。但看了国内外的一些评论后，觉得此书偏向文献综述的类型，适合当作参考手册。亚马逊地址：
http://www.amazon.com/Data-Mining-Concepts-Techniques-Management/dp/0123814790/ref=cm_rdp_product

这是一本从数据库角度阐述数据挖掘的书，主要关注从商业数据库的大量事务数据中寻找有用信息的各种方法。数据库和大数据是贯穿全书的核心。
全书大致可以分成两部分。前一部分重点是数据仓库的构建以及在此过程中的数据整合与化简，对于数据库的设计与数据整理很有启发...

简单来说几句吧。很高兴看到这本书的作者之一Jiawei Han是中国人，先自豪一下。这本书最大的特点就是概念性强（相对于<http://book.douban.com/subject/1820179/>，《数据挖掘中的实用机器学习工具及技术》），从数据仓库到关联规则，从聚类到神经网络，最后几个章节还有数据挖...

这本书是刚上研究生的时候开始看的，这本书介绍的数据挖掘基本上是从数据库的概念出发的，对各种算法都有提及，但是很多算法基本上是语焉不详，对于刚开始学习数据挖掘和机器学习的学生来说，能对数据挖掘的基本概念有所了解，对算法也只能了解个大概了。如果不是纯搞数据仓库...

讲的很不错，就死实现起来有点麻烦。不知道apriori算法大家怎么实现的？主要是采用什么数据结构存储。

开阔眼界非常好 本科的基础不扎实的建议skip这本书吧 Data Mining 可是硕士博士们做的事情

[数据挖掘 下载链接1](#)