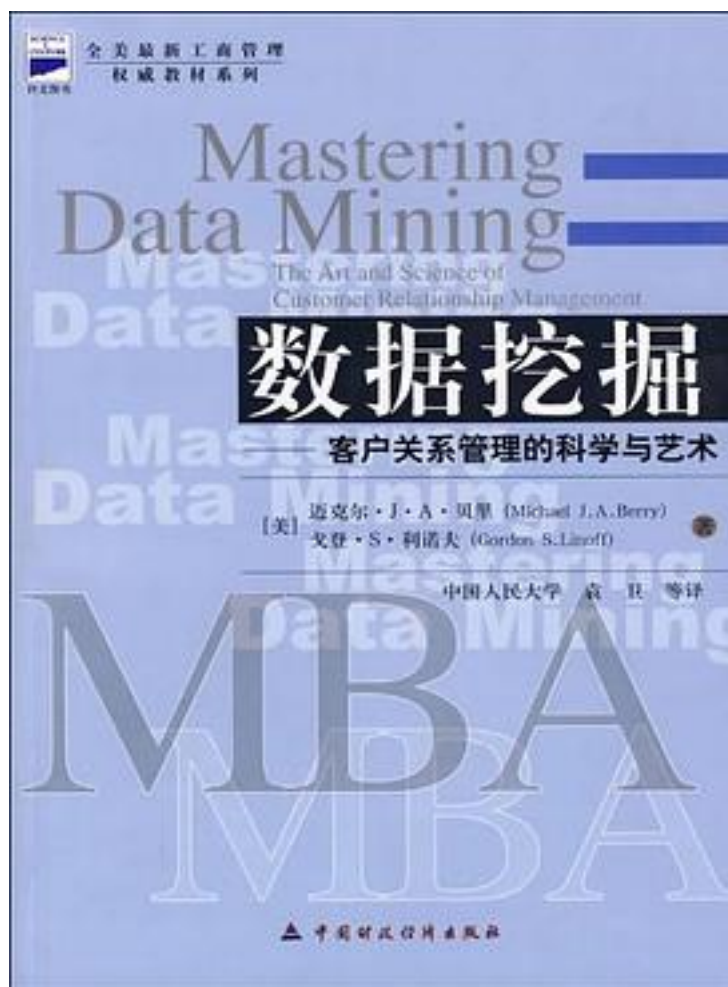


# 数据挖掘



[数据挖掘\\_下载链接1](#)

著者:Jiawei Han

出版者:高等教育出版社

出版时间:2006年01月

装帧:平装

isbn:9787040100419

《数据挖掘:概念与技术(影印版)》是由高等教育出版社出版。

作者介绍:

Jiawei Han(韩家炜)，是伊利诺伊大学厄巴纳-尚佩恩分校计算机科学系的Bliss教授。他因知识发现和数据挖掘研究方面的贡献而获得许多奖励，包括ACM SIGKDD创新奖(2004)、IEEE计算机学会技术成就奖(2005)和IEEE W.Wallace McDowell奖(2009)。他是ACM和IEEE会士。他还担任《ACM Transactions on Knowledge Discovery from Data》的执行主编(2006—2011)和许多杂志的编委，包括《IEEE Transactions on Knowledge and Data Engineering》和《Data Mining Knowledge Discovery》。

拥有加拿大康考迪亚大学计算机科学硕士学位，现在加拿大西蒙弗雷泽大学从事博士后研究工作。

目录:

[数据挖掘\\_下载链接1](#)

标签

数据挖掘

计算机

已入柜

G

CS

评论

为了能在豆瓣网胜任数据分析师的工作，工作开始之前的国庆节在家看了7天书。

-----

-----  
2004年这个概念正热，当作畅销书买的，此后一直落灰。2015年前后拿出来开始阅读，由于过于枯燥无味，且英文阅读困难，直到今年才最终读完。  
这本书我虽然没读太明白，但我一字一句都认真读了，从字里行间可以感受到作者确实认真写了，作者也是这方面的行家里手。  
另外需要说明的是，这本书讲的是理论，不要指望通过这本书速成。就像学完了大学电路课程，也不会修理电视机一样。

-----  
韩老师的力作，非常好的数据挖掘教材

-----  
[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](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很不错的网站给大家。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。  
但是在这种不面善的情况，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，...

-----  
我擦了个擦，  
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...

-----  
这本书应该是大二前后买的（16-17年？），当时大致读了一遍（主要是算法和模型部分），感觉废话巨多，类似于科普。

研究生这本书又作为数据挖掘课程（2学分）教材使用，跳过了第四章、第五章琐碎/过时的数据仓库、数据立方体技术，以及第七章高级模式挖掘和第十一章高级聚类分...

-----  
-----  
-----  
-----  
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。浙大面试时问了我一个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](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。

-----  
[数据挖掘\\_下载链接1](#)