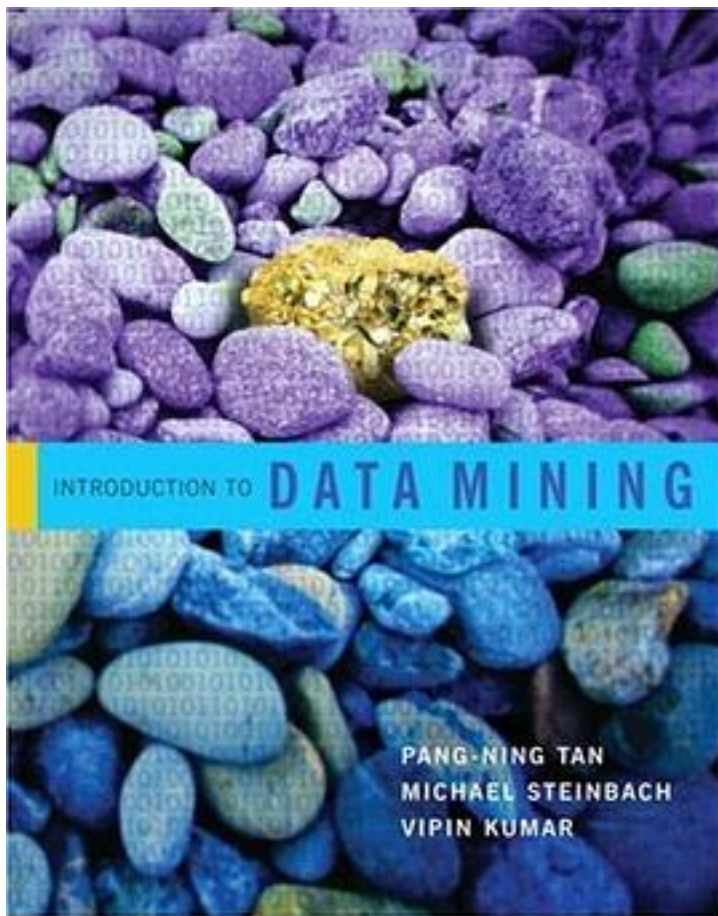


Introduction to Data Mining



[Introduction to Data Mining_下载链接1_](#)

著者:Pang-Ning Tan

出版者:Pearson

出版时间:2013-7-17

装帧:Paperback

isbn:9781292026152

Introduction

Rapid advances in data collection and storage technology have enabled or

ganizations to accumulate vast amounts of data. However, extracting useful information has proven extremely challenging. Often, traditional data analysis tools and techniques cannot be used because of the massive size of a data set. Sometimes, the non-traditional nature of the data means that traditional approaches cannot be applied even if the data set is relatively small. In other situations, the questions that need to be answered cannot be addressed using existing data analysis techniques, and thus, new methods need to be developed.

Data mining is a technology that blends traditional data analysis methods with sophisticated algorithms for processing large volumes of data. It has also opened up exciting opportunities for exploring and analyzing new types of data and for analyzing old types of data in new ways. In this introductory chapter, we present an overview of data mining and outline the key topics to be covered in this book. We start with a description of some well-known applications that require new techniques for data analysis.

Business Point-of-sale data collection (bar code scanners, radio frequency identification (RFID), and smart card technology) have allowed retailers to collect up-to-the-minute data about customer purchases at the checkout counters of their stores. Retailers can utilize this information, along with other business-critical data such as Web logs from e-commerce Web sites and customer service records from call centers, to help them better understand the needs of their customers and make more informed business decisions.

Data mining techniques can be used to support a wide range of business intelligence applications such as customer profiling, targeted marketing, workflow management, store layout, and fraud detection. It can also help retailers

作者介绍:

Pang-Ning

Tan现为密歇根州立大学计算机与工程系助理教授，主要教授数据挖掘、数据库系统等课程。此前，他曾是明尼苏达大学美国陆军高性能计算研究中心副研究员（2002-2003）。

Michael Steinbach 明尼苏达大学计算机与工程系研究员，在读博士。

Vipin

Kumar明尼苏达大学计算机科学与工程系主任，曾任美国陆军高性能计算研究中心主任。他拥有马里兰大学博士学位，是数据挖掘和高性能计算方面的国际权威，IEEE会士。

目录: Chapter 1. Introduction

Pang-Ning Tan/Michael Steinbach/Vipin Kumar 1

Chapter 2. Data

Pang-Ning Tan/Michael Steinbach/Vipin Kumar 19

Chapter 3. Exploring Data

Pang-Ning Tan/Michael Steinbach/Vipin Kumar 97

Chapter 4. Classification: Basic Concepts, Decision Trees, and Model Evaluation

Pang-Ning Tan/Michael Steinbach/Vipin Kumar 145

Chapter 5. Classification: Alternative Techniques

Pang-Ning Tan/Michael Steinbach/Vipin Kumar 207

Chapter 6. Association Analysis: Basic Concepts and Algorithms

Pang-Ning Tan/Michael Steinbach/Vipin Kumar 327

Chapter 7. Association Analysis: Advanced Concepts

Pang-Ning Tan/Michael Steinbach/Vipin Kumar 415

Chapter 8. Cluster Analysis: Basic Concepts and Algorithms

Pang-Ning Tan/Michael Steinbach/Vipin Kumar 487

Chapter 9. Cluster Analysis: Additional Issues and Algorithms

Pang-Ning Tan/Michael Steinbach/Vipin Kumar 569

Chapter 10. Anomaly Detection

Pang-Ning Tan/Michael Steinbach/Vipin Kumar 651

Appendix B: Dimensionality Reduction

Pang-Ning Tan/Michael Steinbach/Vipin Kumar 685

Appendix D: Regression

Pang-Ning Tan/Michael Steinbach/Vipin Kumar 703

Appendix E: Optimization

Pang-Ning Tan/Michael Steinbach/Vipin Kumar 713

Copyright Permissions

Pang-Ning Tan/Michael Steinbach/Vipin Kumar 724

Index 725

• • • • • ([收起](#))

[Introduction to Data Mining 下载链接1](#)

标签

数据挖掘

mining

data

DataMining

评论

挺容易的

[Introduction to Data Mining 下载链接1](#)

书评

我的习惯就是在蹲坑的时候读一些艰涩高深的科学读物，这样有助于我在排泄的时候大脑保持高度的兴奋状态，不至于被熏晕或者不至于被引人入胜的小说情节所陶醉最后导致肛痿……但是，这本书另我惊诧了……
第一他不艰涩，是我读到过的关于统计、关于数据、关于计算的最科普的读...

该书特点：以实例为重，给出了常用算法的伪代码，和《模式识别》、《模式分类》等专著比起来，该书略去了各个定理的证明部分，并通过大量枚举具体的分类实例，来简要说明算法的流程和意义。
根据个人的体验，觉得这本书作为第一本数据挖掘的入门读物是再恰当不过的了。...

这本书介绍的比较全面，某些内容在一般的书中是很少介绍的，内容浅显易懂。本人开始看中文版的，觉的中文版的写的不错，后来又看英文版的，就发现中文版的差太多了，推荐英文版的

我是拿这本书当作课程书的，这本书基本上涵盖了数据挖掘的许多经典算法，分类，聚类，关联规则。比较适合对数据挖掘感兴趣的人，这本书看完之后基本上就可以进行对数据的分析，挖掘了。然而这仅仅是一门入门书，对于理论部分并没有做过多的解释。如果想进一步的了解理论知识， ...

屎一样狗屁不通的翻译。原文：As a result, Z is as likely to be chosen for splitting as the interacting but useful attributes, X and Y. 译文：因此，Z 可能被选作划分有相互作用但有效的属性 X 和 Y。
还有其他很多地方就不一一列举了，本来作为入门读物，很多东西就...

作为数据挖掘导论，这本书基本上已经做到了。书中介绍了很多数据挖掘方面相关的概念和方法，对于入门来讲是很友好的。因为刚刚看完机器学习的书，所以前半部分基本不需要看了。后面的关联分析和聚类方法还是可以一看的。虽然这本书没有实际操作的内容，但是让人大概了解了数据...

主要是一些理论的讲解，对数据挖掘的总体起一个概述的作用，偏向于实际应用的较少！对各种算法也只是简单进行说明，然后进行应用，对于刚刚接触数据挖掘的同学有一些意义 内容涵盖方方面面，对于要深挖某个主题的话需要另找书籍结合阅读

看我截图吧 <http://weibo.com/1677386655/zu8O4ci9O> therefore, if we compute the k-dist for all the data points for some k, sort them in increasing order, and then plot the sorted values, we expect to see a sharp change at the value of k-dist that correspon...

我是非数据挖掘领域，想了解数据挖掘领域的知识，但这本书还是有点太专业，太多的知识和算法看不懂，只是浏览了一下概念性的知识
有没有介绍更通俗的数据挖掘的书，或者注重方法不注重算法的书，希望能有高人指点
——

统计学经典入门书籍，对数据处理、分类、相关分析、聚类等方面做了事无巨细的讲解，兼顾通俗性和理论推导，浏览一遍目录就会发现，这不就是机器学习嘛！
看这书名一开始以为这只是一本讲数据抓取、数据分析的书籍，这比市面上一些夸夸其谈机器学习、人工智能的书要低调很多，而...

为什么没有探讨回归问题？为什么没有探讨回归问题？为什么没有探讨回归问题？为什么没有探讨回归问题？为什么没有探讨回归问题？为什么没有探讨回归问题？为什么没有探讨回归问题？为什么没有探讨回归问题？为什么没有探讨回归问题？为什么没有探讨回...

The book is used as a textbook for my data mining class. It covers all fundamental theories and concepts of data mining, and it explained everything in a quite easy-to-understand and detailed manner. It is suggested to have a good comprehension of some math...

Chapter2 和 Chapter3
一大堆废话，基本都是初中高中教的！！！好像跳过这些章节！！！ Chapter2 和 Chapter3 一大堆废话，基本都是初中高中教的！！！好像跳过这些章节！！！
Chapter2 和 Chapter3
一大堆废话，基本都是初中高中教的！！！好像跳过这些章节！！！！

它是我关于数据挖掘这一方向的入门书。
书中讲了很多基础的数据挖掘算法，读完以后可以对这些算法的基本思想有个了解。书中的例子也很详尽，还是不错的。
但是研究生期间是指望发论文的，这些算法从学术上来说，只能算基础入门了。至于它们在实际工业应...

这本书写得逻辑性比较强，全面，而且我觉得涉及的东西也比较底层，让我们了解一些算法的基本原理是非常重要的。如果，网上的机器学习相关文章看不懂的话，可以从这本书入手。中文版的只看过一点点，感觉完全没逻辑性，完全没感觉。翻译出来完全就变味了，毕竟是语言习惯上的...

给出了DataMining的一般性解决思路，全面易懂，很适合给初学者扫盲。加之与原版大概400+RMB比较起来，不禁觉得还是祖国好哇。。。PS:据说巴基斯坦卖得更便宜。。。

[Introduction to Data Mining 下载链接1](#)