

《数据挖掘算法与Clementine实践》

图书基本信息

书名：《数据挖掘算法与Clementine实践》

13位ISBN编号：9787302235019

10位ISBN编号：7302235015

出版时间：2011-4

出版社：清华大学

作者：熊平

页数：237

版权说明：本站所提供下载的PDF图书仅提供预览和简介以及在线试读，请支持正版图书。

更多资源请访问：www.tushu111.com

《数据挖掘算法与Clementine实践》

内容概要

《数据挖掘算法与Clementine实践》主要介绍了几种最成熟的数据挖掘方法，并针对每种方法，介绍了应用最广泛的几种实现算法。书中以Clementine12.0为平台，用实例介绍了每种算法的具体应用。全书各章分别介绍了数据挖掘和Clementine软件、决策树分类方法(包括ID3、C4.5、C5.0、CART等算法)、聚类分析方法(包括K-Means算法和TwoStep算法)、关联规则挖掘方法(包括Apriori算法、CARMA算法和序列模式挖掘算法)、数据筛选算法(包括特征选择算法和异常检测算法)、回归分析方法(包括线性回归算法和二项Logistic回归)、神经网络构建方法(包括多层感知器网络、RBF网络以及Kohonen网络的构建算法)、时间序列分析方法(包括指数平滑法和ARIMA模型构建方法)。

书籍目录

第1章 数据挖掘概述

- 1.1 数据挖掘简介
- 1.2 数据挖掘过程
- 1.3 数据挖掘方法
- 1.4 数据挖掘工具及软件

第2章 clementine概述

- 2.1 clementine简介
- 2.2 clementine基本操作

第3章 决策树

- 3.1 分类与决策树概述
- 3.2 id3、c4.5与c5.0
- 3.3 cart

第4章 聚类分析

- 4.1 聚类分析概述
- 4.2 k-means算法
- 4.3 twostep算法

第5章 关联规则

- 5.1 关联规则概述
- 5.2 apriori算法
- 5.3 carma算法
- 5.4 序列模式

第6章 数据筛选

- 6.1 特征选择
- 6.2 异常检测

第7章 统计模型

- 7.1 线性回归
- 7.2 项logistic回归

第8章 神经网络

- 8.1 神经网络原理
- 8.2 多层感知器与rbf网络
- 8.3 kohonen网络

第9章 时间序列分析与预测

- 9.1 时间序列概述
- 9.2 指数平滑法
- 9.3 arima模型

章节摘录

版权页：插图：统计方法COBWEB-是一个常用的且简单的增量式概念聚类方法。它的输入对象是采用符号量（属性一值）对来加以描述的。采用分类树的形式来创建一个层次聚类。CLASSIT是COBWEB的另一个版本。它可以对连续取值属性进行增量式聚类。它为每个节点中的每个属性保存相应的连续正态分布（均值与方差）；并利用一个改进的分类能力描述方法，即不像COBWEB那样计算离散属性（取值）和而是对连续属性求积分。但是CLASSIT方法也存在与COBWEB类似的问题。因此它们都不适合对大数据库进行聚类处理。传统的聚类算法已经比较成功地解决了低维数据的聚类问题。但是由于实际应用中数据的复杂性，在处理许多问题时，现有的算法经常失效，特别是对于高维数据和大型数据的情况。因为传统聚类方法在高维数据集中进行聚类时，主要遇到两个问题。首先，高维数据集中存在大量无关的属性使得在所有维中存在簇的可能性几乎为零。另外，高维空间中数据较低维空间中数据分布要稀疏，其中数据间距离几乎相等是普遍现象，而传统聚类方法是基于距离进行聚类的，因此在高维空间中无法基于距离来构建簇。高维聚类分析已成为聚类分析的一个重要研究方向。同时高维数据聚类也是聚类技术的难点。随着技术的进步使得数据收集变得越来越容易，导致数据库规模越来越大、复杂性越来越高，如各种类型的贸易交易数据、Web文档、基因表达数据等，它们的维度（属性）通常可以达到成百上千维，甚至更高。但是，受“维度效应”的影响，许多在低维数据空间表现良好的聚类方法运用在高维空间上往往无法获得好的聚类效果。高维数据聚类分析是聚类分析中一个非常活跃的领域，同时它也是一个具有挑战性的工作。

《数据挖掘算法与Clementine实践》

编辑推荐

《数据挖掘算法与Clementine实践》是由清华大学出版社出版的。

精彩短评

- 1、操作讲解比较详细，结果分析也还行。只是案例相对简单，缺乏深度
- 2、不错的指导书，每个操作步骤都写的很清楚~
- 3、有例子，实用，没那么虚
- 4、大概翻阅一下，感觉还行
- 5、还行吧，好久没看到了，被老师拿走就没还我
- 6、比较简单，入门可以看看，写得很浅显。
- 7、内容尚可，但是有些地方写的不是很详细。
- 8、内容比较新颖，难得有很多实验数据可以提供
- 9、中午刚收到书，内容还没看，但是书的封面很脏，书左上角白色区域有很大一片黑色污渍，黑有一个戳穿的小孔，书左下角的粘合处有很深的挤压凹进去的痕迹，总之对书很不满意，影响心情。由于书是急着用所以没有申请推过。真的怀疑这些书是不是实体店里有瑕疵销售不出去的才拿到网上用相对便宜价格处理掉。
- 10、因为之前一直用SPSS做数据分析，现在计划用Clementine做一下数据挖掘。所以买了这本书读了一下，内容介绍的比较清晰和易于理解。不足之处是书太薄了，不够详细（比如很多指标的含义）。作为一本入门书，我觉得足够了。
- 11、纸张一般,正在看
- 12、软件的介绍不多 作为打破学习Clementine的心里障碍还是很有帮助的
- 13、配送速度很OK，书的内容页基本满意，围绕几个主要的数据挖掘算法以clementine为工具简单介绍了应用，但是感觉内容太简单，书的质量比较一般
- 14、内容感觉很一般，有点小失望！
- 15、速度较快，比较满意！
- 16、没有全部看完，但是感觉还可以。
- 17、入门教材，物有所值，值得初级人员入手
- 18、看起来不错，例子挺多
- 19、还算可以，就是纸质太差，也有点失望
- 20、写的一般
- 21、案例偏简单，适合快速入门。
- 22、没来得及看，总体还行。
- 23、基础且实用,书不错,可以用于基础的学习
- 24、could be better
- 25、这本书主要是讲如何使用cleantime工具的，算法讲解一般
- 26、讲解使用Clementine工具进行数据挖掘的算法。很好
- 27、老公评价挺高的。
- 28、虽然书中很多内容出自Clementine12.0的用户手册，但是很有针对性。书中的例子也很详细，算法和软件上手的内容很适合入门者。值得推荐。
- 29、书的内容跟网上的差不多，是正版！

《数据挖掘算法与Clementine实践》

精彩书评

1、因为之前一直用SPSS做数据分析，现在计划用Clementine做一下数据挖掘。所以买了这本书读了一下，内容介绍的比较清晰和易于理解。不足之处是书太薄了，不够详细（比如很多指标的含义）。作为一本入门书，我觉得足够了。

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:www.tushu111.com