

<<数据挖掘技术及其应用>>

图书基本信息

书名：<<数据挖掘技术及其应用>>

13位ISBN编号：9787313066107

10位ISBN编号：7313066104

出版时间：2011-1

出版时间：上海交通大学出版社

作者：杨杰,姚莉秀

页数：263

字数：319000

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>

<<数据挖掘技术及其应用>>

内容概要

《数据挖掘技术及其应用(精)》(作者杨杰、姚莉秀)系统地讲述了数据挖掘的基本概念和基本原理,并列举了在相应领域具有参考价值的算法及其改进和应用,是作者多年来从事教学和科研实践的成果。

全书共9章,主要内容有:数据挖掘的基本概念和原理,数据预处理,各种分类、聚类和关联规则提取算法,以及在生物信息学、材料学中的实际应用案例。

《数据挖掘技术及其应用(精)》可用作计算机专业本科高年级学生或研究生的教材或参考书,也可供从事计算机信息处理、数据挖掘、工业优化等有关方面工作的科技人员参考。

<<数据挖掘技术及其应用>>

作者简介

杨杰，男，1964年8月出生于上海，汉族。

获德国汉堡大学计算机系工学博士学位，上海交通大学图像处理与模式识别研究所教授，博士生导师。

姚莉秀，现为上海交通大学图象处理及模式识别研究所副教授、博士。

杭州大学本科毕业，中科院冶金所博士毕业。

<<数据挖掘技术及其应用>>

书籍目录

第1章 导论

- 1.1 数据挖掘技术的源起与发展
- 1.2 数据挖掘的概念
- 1.3 数据挖掘的过程
- 1.4 数据挖掘的功能
- 1.5 数据挖掘的典型应用领域
- 1.6 目前国际上流行的数据挖掘软件

参考文献

第2章 数据预处理

- 2.1 数据清理
- 2.2 数据集成
- 2.3 数据转换
- 2.4 数据约简

参考文献

第3章 维约简——特征选择与特征提取

- 3.1 特征选择
- 3.2 特征提取
- 3.3 基于谱分析的降维框架

参考文献

第4章 关联规则提取与粗糙集

- 4.1 基本概念
- 4.2 经典的关联规则挖掘算法
- 4.3 模糊关联规则的发现
- 4.4 数量属性关联规则的挖掘
- 4.5 面向不确定知识的关联规则挖掘——粗糙集理论与应用
- 4.6 基于粗糙集和微粒群算法的特征选择(PSORSFS)
- 4.7 基于有序PSO的粗糙集近似熵约简
- 4.8 基于模糊粗糙集的最近邻聚类分类算法

参考文献

第5章 分类原理与方法

- 5.1 一般概念
- 5.2 基于归纳的传统决策树方法
- 5.3 超平面决策树方法
- 5.4 复合式评价函数
- 5.5 模糊类别的决策树方法
- 5.6 基于模糊极小极大网络的模糊规则提取与分类
- 5.7 Linear Map(LMAP)方法与包容型数据

参考文献

第6章 统计学习理论与支持向量机

- 6.1 简介
- 6.2 统计学习理论的主要内容
- 6.3 支持向量机理论
- 6.4 基于测地距离的SVM分类算法
- 6.5 基于SOR(Successive Over Relaxation)的支持向量回归训练方法

<<数据挖掘技术及其应用>>

参考文献

第7章 聚类分析

7.1 聚类的基本概念

7.2 常见聚类算法

7.3 特征空间属性加权模糊核聚类算法

7.4 基于信息理论的合作模糊聚类算法研究

7.5 基于密度和网格的子空间聚类算法

参考文献

第8章 数据挖掘在生物信息学中的应用

8.1 基于集成分类器的蛋白序列分析

8.2 聚类分析在基因表达数据中的应用

8.3 基于有监督聚类算法的蛋白三维结构分类

参考文献

第9章 数据挖掘在合金相图研究中的应用

9.1 国内外相图研究现状

9.2 相图研究的原子参数—数据挖掘方法

9.3 研究三元合金系中间化合物形成规律的原理与方法

9.4 国内外相图研究现状：三元合金系中间化合物形成规律研究

参考文献

<<数据挖掘技术及其应用>>

章节摘录

版权页：插图：1.4 数据挖掘的功能数据挖掘通过对现有数据记录的分析，预测未来趋势及行为，做出基于知识的决策。

数据挖掘的目标是从大量数据中发现隐含的、有意义的知识，主要有以下几类功能。

1) 概念描述被分析的数据称为目标数据集，像程序设计语言C++里面的对象与类一样，对于一个数据集，我们可以通过其共有的属性和行为来描述它，最终获得简明、准确的描述，使之更具代表性。

2) 相关分析（关联分析）就是从给定的数据集发现频繁出现的项集模式知识，即发现各属性之间的关联关系并用关联规则描述出来。

3) 分类和回归预报根据一系列已知数据，训练产生一套能描述或区别数据的类别或概念的模型，并能够根据这个模型来预测未知数据的结果。

人脸识别、指纹识别、商业中的客户识别分类、工业上故障诊断等都是分类问题。

4) 聚类根据物以类聚原则，利用属性特征将数据集合分成为由类似的数据组成的多个类的过程称为聚类。

聚类后，同一类之间的数据具有很强的相似性而非同类之间的数据具有很强的非相似性。

5) 趋势分析实现前面提到的四种功能时，事件产生的顺序信息都被忽略，被简单地作为一条静态的记录来对待。

而趋势分析是对随时间变化的数据对象的变化规律和趋势进行建模描述，根据前一段时期的运动预测下一个时间点的状态。

解决的问题一般可分为两类：（1）总结数据的序列或者变化趋势，如预测股票 / 期货交易，网页点击顺序记录等。

<<数据挖掘技术及其应用>>

编辑推荐

《数据挖掘技术及其应用》是由上海交通大学出版社出版的。

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>