

<<数据挖掘理论与技术>>

图书基本信息

书名：<<数据挖掘理论与技术>>

13位ISBN编号：9787121189890

10位ISBN编号：7121189895

出版时间：2013-1

出版时间：电子工业出版社

作者：罗森林

页数：242

字数：400000

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>

<<数据挖掘理论与技术>>

内容概要

本书梳理了数据挖掘理论与技术的知识点，注重领域内核心思想、原理、方法的论述及国内外最新研究进展的融入，内容上系统、全面、先进。

全书共9章，主要包括数据挖掘基础知识，概率论与数理统计，数据挖掘效果评价，数据预处理，数据仓库，数据分类分析，数据聚类分析，关联规则发现，统计预测方法等。

在讨论算法的同时引入应用实例，强调应用方法包括算法特点、参数选择、结果评价等方面的分析，理论联系实际，有利于算法的快速掌握和有效运用。

<<数据挖掘理论与技术>>

书籍目录

第1章 绪论

1.1 数据挖掘产生的背景

1.1.1 技术背景

1.1.2 理论基础

1.1.3 数据挖掘相关概念

1.2 数据挖掘知识基础

1.2.1 基本概念及特点

1.2.2 数据集

1.2.3 功能与分类

1.2.4 任务与过程

1.2.5 方法与步骤

1.3 数据挖掘简史与现状

1.3.1 简史

1.3.2 现状

1.4 数据挖掘的技术工具

1.4.1 技术工具

1.4.2 工具选择

1.5 数据挖掘的应用

1.5.1 典型应用

1.5.2 高级应用

1.6 技术难点与发展趋势

1.6.1 常见误解

1.6.2 技术难点

1.6.3 发展趋势

1.7 本章小结

思考题

第2章 概率统计理论基础

2.1 引言

2.2 概率统计知识基础

2.3 随机变量的分布函数

2.3.1 多维随机变量

2.3.2 条件分布

2.4 统计推理

2.5 参数估计

2.5.1 估计理论

2.5.2 最大似然估计

2.5.3 贝叶斯估计

2.6 假设检验

2.7 数据采样方法

2.8 本章小结

思考题

第3章 数据挖掘效果评价

3.1 引言

3.2 模型的评分函数

3.2.1 基本概念

<<数据挖掘理论与技术>>

3.2.2 预测模型的评分函数

3.2.3 描述模型的评分函数

3.3 模型的比较与验证

3.3.1 模型比较

3.3.2 模型验证

3.4 模型的性能提升

3.4.1 增量学习

3.4.2 半监督学习

3.4.3 迁移学习

3.4.4 反模型

3.4.5 Boosting

3.5 模型的建立与使用

3.5.1 模型的建立

3.5.2 模型的理解

3.5.3 模型的使用

3.6 本章小结

思考题

第4章 数据预处理

4.1 引言

4.2 数据预处理知识基础

4.3 数据清理

4.3.1 遗漏值

4.3.2 噪声数据

4.3.3 不一致数据

4.4 数据集成

4.5 数据转换

4.6 数据规约

4.6.1 数据方聚集

4.6.2 维归约

4.6.3 数据压缩

4.6.4 数值归约

4.7 数据离散

4.8 应用实例分析

4.8.1 腹围空缺数值归一化弥补方法

4.8.2 型糖尿病数据预处理

4.9 本章小结

思考题

第5章 数据仓库

5.1 引言

5.2 数据仓库知识基础

5.2.1 基本概念

5.2.2 基本作用

5.2.3 与数据挖掘的关系

5.3 数据仓库中的模型

5.3.1 概念模型

5.3.2 物理模型

5.3.3 元数据模型

<<数据挖掘理论与技术>>

5.3.4 多维数据模型

5.4 数据仓库系统结构

5.4.1 组成

5.4.2 数据仓库概念结构

5.4.3 数据仓库结构类型

5.5 OLAP分析

5.5.1 知识基础

5.5.2 多维分析

5.5.3 OLAP结构

5.5.4 多维数据库

5.5.5 关系数据库

5.6 本章小结

思考题

第6章 数据分类分析

6.1 引言

6.2 分类分析知识基础

6.2.1 基本概念

6.2.2 基本作用

6.2.3 评价方法

6.3 主要技术方法及分析

6.4 贝叶斯分类

6.4.1 朴素贝叶斯分类法

6.4.2 贝叶斯网络

6.4.3 动态贝叶斯网络

6.5 基于决策树的算法

6.5.1 基本思想

6.5.2 ID3算法

6.5.3 C4.5算法

6.5.4 SLIQ算法

6.5.5 SPRINT算法

6.6 神经网络与遗传算法

6.6.1 神经网络

6.6.2 遗传算法

6.7 支持向量机

6.8 粗糙集与模糊集

6.8.1 粗糙集

6.8.2 模糊集

6.9 最大熵模型

6.10 应用实例分析

6.10.1 汉语句义类型识别

6.10.2 特定音频事件识别

6.11 本章小结

思考题

第7章 数据聚类分析

7.1 引言

7.2 聚类分析知识基础

7.2.1 基本概念

<<数据挖掘理论与技术>>

- 7.2.2 基本作用
- 7.2.3 近邻测度
- 7.2.4 评价方法
- 7.3 主要技术方法及分析
- 7.4 基于划分的算法
 - 7.4.1 基本思想
 - 7.4.2 K-means算法
 - 7.4.3 K-medoids算法
 - 7.4.4 CLARANS算法
- 7.5 基于层次的算法
 - 7.5.1 基本思想
 - 7.5.2 BIRCH算法
 - 7.5.3 CURE算法
 - 7.5.4 ROCK算法
 - 7.5.5 Chameleon算法
- 7.6 基于密度的算法
 - 7.6.1 基本思想
 - 7.6.2 DBSCAN算法
 - 7.6.3 OPTICS算法
 - 7.6.4 DENCLUE算法
- 7.7 基于网格的算法
 - 7.7.1 基本思想
 - 7.7.2 STING算法
 - 7.7.3 Wave Cluster算法
 - 7.7.4 CLIQUE算法
- 7.8 基于模型的算法
 - 7.8.1 基本思想
 - 7.8.2 EM算法
 - 7.8.3 COBWEB算法
 - 7.8.4 自组织神经网络
- 7.9 应用实例分析
 - 7.9.1 镜头聚类
 - 7.9.2 文本聚类
- 7.10 本章小结
- 思考题
- 第8章 关联规则发现
 - 8.1 引言
 - 8.2 关联规则发现知识基础
 - 8.2.1 基本概念
 - 8.2.2 评价方法
 - 8.2.3 注意事项
 - 8.3 主要技术方法及分析
 - 8.4 关联规则的基本算法
 - 8.4.1 Apriori算法
 - 8.4.2 FP-树频集算法
 - 8.4.3 CloSpan
 - 8.5 并行和分布式关联规则算法

<<数据挖掘理论与技术>>

- 8.5.1 并行关联规则
- 8.5.2 分布式关联规则
- 8.6 多层次关联规则算法
- 8.7 数量关联规则算法
- 8.8 应用实例分析——蠕虫检测
- 8.9 本章小结
- 思考题
- 第9章 统计预测方法
- 9.1 引言
- 9.2 统计预测方法知识基础
- 9.3 主要技术方法及分析
- 9.4 回归预测方法
- 9.4.1 线性和多元回归
- 9.4.2 非线性回归
- 9.5 Box-Jenkins回归
- 9.6 隐马模型
- 9.6.1 隐马尔可夫模型
- 9.6.2 隐半马尔可夫模型
- 9.7 应用实例分析
- 9.7.1 型糖尿病发病危险状态预测
- 9.7.2 关键人物判定
- 9.8 本章小结
- 思考题
- 参考文献

<<数据挖掘理论与技术>>

编辑推荐

罗森林编著的《数据挖掘理论与技术》梳理了数据挖掘理论与技术的知识点，注重领域内核心思想、原理、方法的论述及国内外最新研究进展的融入，内容上系统、全面、先进。本书可供计算机科学与技术、生命信息工程、软件工程、通信与信息系统等相关学科、专业的学生作为教材或参考书，同时也可供科研人员参考和感兴趣者自学使用。

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>