

<<数据挖掘与知识发现>>

图书基本信息

书名：<<数据挖掘与知识发现>>

13位ISBN编号：9787040304787

10位ISBN编号：7040304783

出版时间：2010-7

出版时间：高等教育出版社

作者：李雄飞等著

页数：303

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>

## <<数据挖掘与知识发现>>

### 内容概要

数据挖掘与知识发现领域中的一些基本理论、研究方法和技术标准，简单介绍了相关产品和工具，讨论了KDD与数据挖掘的概念、数据挖掘对象、知识发现过程、研究方法以及相关的研究领域和应用范围。

第2章～第9章详细地介绍了关联规则、聚类分析、决策树、贝叶斯网络、人工神经网络、支持向量机、粗糙集、模糊集等数据挖掘模型与算法。

第10章讨论了模型选择与模型评估方法。

第11章和第12章简单介绍了数据预处理方法和数据挖掘技术标准、数据挖掘可视化技术和数据挖掘工具开发方法，并简单介绍了数据挖掘产品和工具。

《数据挖掘与知识发现（第2版）》可以作为计算机专业、信息类专业、管理类专业高年级本科生及研究生的教材或参考书，也可供有关人员学习参考。

## 书籍目录

第1章 绪论1.1 引言1.2 KDD与数据挖掘1.2.1 KDD定义1.2.2 KDD过程1.2.3 数据库技术与数据挖掘1.3 数据挖掘的对象与环境1.3.1 数据与系统特征1.3.2 数据结构1.3.3 数据库系统1.4 数据挖掘方法与相关领域1.4.1 数据挖掘相关领域1.4.2 粗糙集1.4.3 聚类1.4.4 关联规则1.4.5 决策树1.4.6 模糊集1.4.7 规则归纳1.4.8 进化计算1.5 KDD系统与应用本章小结习题1第2章 关联规则2.1 引言2.2 关联规则基本模型2.2.1 关联规则基本模型2.2.2 Apriori算法2.2.3 LIG算法2.2.4 FP算法2.3 多级关联规则与多维关联规则2.3.1 多级关联规则2.3.2 多维关联规则2.4 关联规则价值衡量与发展2.4.1 规则价值衡量2.4.2 基于约束的关联规则2.4.3 关联规则新进展本章小结习题2第3章 聚类分析3.1 聚类分析简介3.2 聚类分析中的数据类型3.3 划分方法3.3.1 加均值算法3.3.2 k-中心点算法3.3.3 EM算法3.4 层次方法3.4.1 凝聚的和分裂的层次聚类3.4.2 利用层次方法进行平衡迭代归约和聚类3.4.3 利用代表点聚类3.4.4 采用动态建模技术的层次聚类算法3.5 基于密度的方法3.6 基于网格的方法3.7 基于模型的聚类方法3.8 孤立点分析本章小结习题3第4章 决策树4.1 引言4.2 信息论4.2.1 信息传输与数据挖掘4.2.2 信息论主要概念4.3 ID3算法4.4 决策树的剪枝4.4.1 预剪枝4.4.2 后剪枝4.4.3 决策树的性能评价4.5 决策树算法的改进4.5.1 二叉树决策算法4.5.2 按增益比率估值的方法4.5.3 按分类信息估值的方法4.5.4 按划分距离估值的方法4.6 C4.5 算法4.7 CART算法4.8 SLIQ算法4.9 决策树与数据预处理4.9.1 数据概化与约简4.9.2 抽样方法4.9.3 维归约及特征子集的选取4.9.4 冗余特征子集删除4.9.5 离散化处理4.9.6 改变数据结构4.10 算法改进4.10.1 多决策树综合技术4.10.2 决策树的增量学习本章小结习题4第5章 贝叶斯网络5.1 贝叶斯网络基本概念5.2 不确定性推理与联合概率分布5.3 贝叶斯网络中的独立关系5.3.1 条件独立5.3.2 有向分离和条件独立5.3.3 因果影响独立5.3.4 环境独立5.4 贝叶斯网络学习5.4.1 结构学习5.4.2 搜索算法5.4.3 基于约束的方法5.4.4 参数学习5.5 贝叶斯网络分类器5.5.1 朴素贝叶斯网络分类器5.5.2 半朴素贝叶斯分类器与选择贝叶斯分类器5.5.3 树增广朴素贝叶斯网络分类器5.5.4 广义朴素贝叶斯网络分类器本章小结习题5第6章 人工神经网络6.1 人工神经元及人工神经网络模型6.1.1 M.P模型6.1.2 人工神经元的形式化描述6.1.3 神经网络的分类6.1.4 人工神经网络的学习方式6.2 前向神经网络6.2.1 感知器6.2.2 多层前向神经网络的BP算法6.2.3 径向基函数神经网络6.3 反馈神经网络6.3.1 前向神经网络与反馈神经网络的比较6.3.2 反馈神经网络模型6.3.3 离散型Hopfield神经网络6.3.4 连续型Hopfield神经网络6.3.5 Boltzmann机6.4 自组织竞争神经网络模型6.5 基于人工神经网络的数据挖掘本章小结习题6第7章 支持向量机7.1 学习机器泛化性能的界7.1.1 VC维7.1.2  $R_n$ 中有向超平面对点的打散7.1.3 VC维和参数个数7.1.4 通过最小化 $h$ 最小化界7.1.5 实例7.1.6 结构风险最小化7.2 线性支持向量机7.2.1 可分情形7.2.2 Karush-Kuim-TUckeai 条件7.2.3 测试7.2.4 非可分情形7.3 非线性支持向量机7.3.1 硬间隔非线性支持向量机7.3.2 软间隔非线性支持向量机7.3.3  $\nu$ -SVM分类器7.3.4 处理不平衡数据的加权SVM7.3.5 多类别SVM分类7.3.6 Mercer条件及Mercer定理7.3.7 非线性支持向量机实例7.4 支持向量机的VC维7.5 支持向量机应用7.5.1 手写体数字识别7.5.2 文本分类7.5.3 生物信息学中的SVM应用本章小结习题7第8章 粗糙集8.1 近似空间8.1.1 近似空间与不可分辨关系8.1.2 知识与知识库8.2 近似与粗糙集8.2.1 近似与粗糙集的基本概念8.2.2 粗糙集的基本性质8.3 粗糙集的特征描述8.3.1 近似精度8.3.2 粗糙集隶属函数8.3.3 拓扑特征8.4 知识约简8.4.1 约简与核8.4.2 相对约简和相对核8.5 知识的依赖性8.6 信息系统8.6.1 信息系统的定义8.6.2 分辨矩阵与分辨函数8.7 决策表8.8 决策规则8.9 扩展的粗糙集模型8.9.1 可变精度粗糙集模型8.9.2 相似模型本章小结习题8第9章 模糊集9.1 模糊集定义与隶属函数9.1.1 模糊集定义与隶属函数9.1.2 模糊集合的表示法9.2 模糊集的基本运算9.3 分解定理与扩展原理9.4 模糊集的特征9.5 模糊集的度量9.5.1 模糊度9.5.2 模糊集间的距离9.5.3 模糊集的贴度9.6 模糊关系9.6.1 模糊关系定义9.6.2 模糊关系的运算与性质9.6.3 模糊等价关系与模糊相似关系9.7 模糊聚类分析9.7.1 模糊划分9.7.2 模糊相似系数的标定方法9.7.3 模糊聚类分析9.7.4 传递闭包法9.7.5 最大树法9.7.6 模糊c-均值聚类9.8 模糊集与粗糙集本章小结习题9第10章 模型选择与模型评估10.1 模型的过拟合10.2 没有天生优越的分类器10.3 模型、模型选择和模型评估10.4 简单划分和交叉验证10.5 自助法10.6 Occam剃刀10.7 最小描述长度准则10.8 信息准则10.8.1 Akaike信息准则10.8.2 Bayesian信息准则10.9 比较分类器的方法10.9.1 估计准确率的置信区间10.9.2 比较两个模型的性能10.9.3 比较两种分类法的性能10.10 聚类评估10.10.1 假设检验10.10.2 聚类评估中的假设检验10.10.3 相对准则本章小结习题10第11章 数据预处理与可视化技术11.1 数据清理11.1.1 填补空缺值11.1.2 消除噪声数

据11.1.3 实现数据一致性11.2 数据集成与转换11.2.1 数据集成11.2.2 数据转换11.3 数据归约与浓缩11.3.1 数据立方体聚集11.3.2 维归约11.3.3 数据压缩11.3.4 数值归约11.4 概念分层11.4.1 概念分层的概念11.4.2 概念分层的类型11.4.3 数值数据的概念分层与离散化11.4.4 分类数据的概念分层11.5 可视化技术概述11.5.1 可视化技术分类11.5.2 可视化技术在数据挖掘中的应用11.6 过程可视化11.7 数据可视化11.7.1 折线图11.7.2 复合饼图11.7.3 散点图11.7.4 盒图11.7.5 平行坐标法11.7.6 圆环分段表示11.8 结果可视化11.8.1 关联规则11.8.2 分类11.8.3 聚类本章小结习题11第12章 数据挖掘工具与产品12.1 数据挖掘标准12.1.1 数据挖掘标准化概述12.1.2 数据挖掘过程标准12.1.3 数据挖掘接口标准12.1.4 数据挖掘的语言标准12.1.5 数据挖掘的Web标准12.1.6 数据挖掘标准的应用与未来发展趋势12.2 数据挖掘工具的介绍12.2.1 数据挖掘工具发展过程概述12.2.2 数据挖掘工具简介12.2.3 WEKA12.2.4 SPSS12.3 数据挖掘产品分析12.3.1 通用数据挖掘产品12.3.2 专用挖掘产品本章小结习题12附录中英文术语对照参考文献

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>