

<<数据仓库与数据挖掘>>

图书基本信息

书名：<<数据仓库与数据挖掘>>

13位ISBN编号：9787302106883

10位ISBN编号：7302106886

出版时间：2005-6-1

出版时间：清华大学出版社

作者：安淑芝

页数：206

字数：323000

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>

## <<数据仓库与数据挖掘>>

### 内容概要

本书是一本介绍数据仓库和数据挖掘的图书。

全书力求深入浅出、通过浅显易懂的语言及实例介绍数据仓库与数据挖掘的基本概念及相关理论。

从数据仓库的定义、结构、设计、数据访问方法及应用等方面对数据仓库做了较详细的介绍。

从数据挖掘的定义、数据预处理方法、数据挖掘发现知识的类型及数据挖掘常用算法等几方面对数据挖掘的基本知识和算法等理论做了介绍。

本书特别介绍了SQL Server2000数据挖掘工具应用和SPSS数据挖掘工具应用。

最后，给出了一个数据挖掘的应用实例。

本书总的指导思想是在掌握基本知识和基本理论的基础上，更强调实际应用能力的培养。

本书可作为普通高等院校计算机科学与技术专业、软件工程专业或信息类等其他相关专业的教材，也可作为有关数据仓库与数据挖掘方面的培训教材，以及所有想学习数据仓库与数据挖掘知识的人的自学用书。

## &lt;&lt;数据仓库与数据挖掘&gt;&gt;

## 书籍目录

第1章 绪论	1.1 初识数据挖掘	1.1.1 数据挖掘的产生	1.1.2 数据挖掘的应用价值	1.1.3 数据挖掘的发展过程	1.1.4 数据挖掘的定义	1.2 初识数据仓库	1.2.1 数据仓库的产生	1.2.2 数据仓库的应用价值	1.2.3 数据仓库的发展过程	1.2.4 数据仓库的定义	1.2.5 数据仓库与数据挖掘的关系	1.3 进一步理解数据挖掘	1.3.1 数据挖掘的功能	1.3.2 数据挖掘常用技术	1.3.3 数据挖掘的过程	1.4 数据挖掘应用实例	1.4.1 应用领域	1.4.2 典型案例	1.5 数据挖掘的发展趋势																																
	1.5.1 数据挖掘研究方向	1.5.2 数据挖掘应用的热点	小结	习题	第2章 数据仓库	2.1 进一步深入理解数据仓库的定义	2.1.1 数据仓库的数据是面向主题的	2.1.2 数据仓库的数据是集成的	2.1.3 数据仓库的数据是不可更新的	2.1.4 数据仓库的数据是随时间不断变化的	2.2 数据仓库的结构	2.2.1 元数据	2.2.2 粒度的概念	2.2.3 分割问题	2.2.4 数据仓库中的数据组织形式	2.3 数据仓库的说明——标准手册	2.4 数据仓库的清理	2.5 数据仓库系统的设计	2.5.1 数据仓库系统设计方法	2.5.2 数据仓库设计的三级数据模型	2.5.3 提高数据仓库的性能	2.5.4 数据仓库设计步骤	2.6 数据仓库数据的访问	2.6.1 数据仓库数据的直接访问	2.6.2 数据仓库数据的间接访问	2.7 数据仓库的应用	2.7.1 数据仓库的主要应用领域	2.7.2 数据仓库应用实例	小结	习题																					
第3章 数据预处理	3.1 数据预处理的目的是功能	3.1.1 原始数据中存在的问题	3.1.2 数据预处理的方法和功能	3.2 数据清理	3.2.1 处理空缺值	3.2.2 噪声数据的处理	3.3 数据集成和变换	3.3.1 数据集成	3.3.2 数据变换	3.4 数据归约	3.4.1 数据归约的方法	3.4.2 数据立方体聚集	3.4.3 维归约	3.4.4 数据压缩	3.4.5 数值归约	3.4.6 离散化与概念分层生成	小结	习题	第4章 数据挖掘发现知识的类型	4.1 广义知识	4.1.1 广义知识的概念	4.1.2 广义知识的发现方法	4.2 关联知识	4.2.1 关联知识的概念	4.2.2 关联知识的发现方法	4.2.3 关联规则应用实例	4.3 分类知识	4.3.1 分类知识的概念	4.3.2 分类知识的发现方法	4.3.3 分类知识应用实例	4.4 预测型知识	4.4.1 预测型知识的概念	4.4.2 预测型知识的发现方法	4.4.3 预测型知识应用实例	4.5 偏差型知识	4.5.1 偏差型知识的概念	4.5.2 偏差型知识的发现方法	小结	习题												
第5章 数据挖掘中常用算法	5.1 神经网络算法	5.1.1 神经网络的概念	5.1.2 神经网络的计算机模型	5.1.3 定义神经网络拓扑	5.1.4 基于神经网络的算法	5.2 使用候选项集找频繁项集 (Apriori) 算法	5.2.1 关联规则的分类	5.2.2 Apriori算法	5.2.3 从频繁项集产生关联规则	5.3 决策树算法	5.3.1 信息论的基本原理	5.3.2 ID3算法	5.3.3 树剪枝	5.3.4 由决策树提取分类规则	5.4 聚类分析	5.4.1 聚类分析的概念	5.4.2 聚类分析中的数据类型	5.4.3 几种主要的聚类分析方法	5.4.4 聚类分析算法	小结	习题	第6章 数据挖掘的工具及其应用	6.1 SQL Server 2000数据挖掘工具应用	6.1.1 安装要求	6.1.2 安装过程	6.1.3 Analysis Services功能介绍	6.1.4 Analysis Services的优点	6.1.5 创建数据挖掘模型	6.1.6 查看和分析挖掘结果	6.1.7 聚类模型	6.2 SPSS数据挖掘工具应用	6.2.1 安装SPSS Clementine	6.2.2 SPSS Clementine 8.0工作环境介绍	6.2.3 Clementine应用的结构	6.2.4 Clementine的使用	6.2.5 挖掘模型的建立和执行	小结	习题	第7章 数据挖掘应用实例	7.1 实例背景	7.2 决策树算法	7.2.1 数据挖掘中的分类算法	7.2.2 决策树的概念	7.3 实例开发	7.3.1 实例开发前的准备	7.3.2 实例的系统结构	7.3.3 决策树算法模块	7.3.4 算法的程序实现	7.4 核心源程序	小结	参考文献

## <<数据仓库与数据挖掘>>

### 编辑推荐

本书本着“在掌握基本知识和基本理论的基础上，强调实际应用能力的培养”的指导思想，在写作上力求体现如下特点：

- 采用尽可能浅显易懂的语言，循序渐进地表达知识内容；
- 概念和具体的方法、工具相结合，使知识具体化，不枯燥；
- 尽可能结合应用的实例，使理论和实际相结合，达到学以致用效果。

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>