

<<数据仓库与数据挖掘教程>>

图书基本信息

书名：<<数据仓库与数据挖掘教程>>

13位ISBN编号：9787302259138

10位ISBN编号：7302259135

出版时间：2011-11

出版时间：清华大学出版社

作者：陈文伟

页数：304

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教程>>

### 内容概要

数据仓库与数据挖掘是决策支持的两项重要技术，它们共同的特点是都需要利用大量的数据资源，并从数据资源中提取信息和知识。

由于数据资源丰富，因此数据仓库与数据挖掘的决策支持效果显著。

本书系统介绍数据仓库原理，联机分析处理，数据仓库设计与开发，数据仓库的决策支持，数据挖掘原理，基于信息论的决策树方法，基于集合论的粗糙集方法、k—均值聚类、关联规则挖掘，仿生物技术的神经网络，遗传算法，公式发现，知识挖掘，文本挖掘与web挖掘。

本书从数据仓库的兴起来说明决策支持的特点，从数据挖掘的理论基础来说明数据挖掘的方法，并通过实例来详细讲解。

希望读者在学习之后，亲自在计算机上去实践，这样才能更有效地掌握数据挖掘的方法。

<<数据仓库与数据挖掘教程>>

书籍目录

第1章 数据仓库与数据挖掘概述

- 1.1 数据仓库的兴起
- 1.2 数据挖掘的兴起
- 1.3 数据仓库和数据挖掘的结合

习题1

第2章 数据仓库原理

- 2.1 数据仓库结构体系
- 2.2 数据仓库数据模型
- 2.3 数据抽取、转换和装载
- 2.4 元数据

习题2

第3章 联机分析处理

- 3.1 olap概念
- 3.2 olap的数据模型
- 3.3 多维数据的显示
- 3.4 oalp的多维数据分析

习题3

第4章 数据仓库设计与开发

- 4.1 数据仓库分析与设计
- 4.2 数据仓库开发
- 4.3 数据仓库技术与开发的困难

习题4

第5章 数据仓库的决策支持

- 5.1 数据仓库的用户
- 5.2 数据仓库的决策支持与决策支持系统
- 5.3 数据仓库应用实例

习题5

第6章 数据挖掘原理

- 6.1 数据挖掘综述
- 6.2 数据挖掘方法和技术
- 6.3 数据挖掘的知识表示

习题6

第7章 信息论方法

- 7.1 信息论原理
- 7.2 决策树方法
- 7.3 决策规则树方法

习题7

第8章 集合论方法

- 8.1 粗糙集方法
- 8.2 k—均值聚类
- 8.3 关联规则挖掘

习题8

第9章 神经网络

- 9.1 神经网络概念与感知机
- 9.2 反向传播网络

<<数据仓库与数据挖掘教程>>

9.3 径向基函数网络

9.4 神经网络的几何意义

习题9

第10章 遗传算法与进化计算

10.1 遗传算法

10.2 基于遗传算法的分类学习系统

10.3 进化计算

习题10

第11章 公式发现

11.1 公式发现概述

11.2 科学定律重新发现系统

11.3 经验公式发现系统

习题11

第12章 知识挖掘

12.1 变换规则的知识挖掘

12.2 软件进化规律的知识挖掘

习题12

第13章 文本挖掘与web挖掘

13.1 文本挖掘概述

13.2 文本挖掘

13.3 web挖掘

习题13

参考文献

## <<数据仓库与数据挖掘教程>>

### 章节摘录

版权页：插图：2.数据仓库用于决策分析随着决策分析需求的扩大，兴起了支持决策的数据仓库。它是以决策主题需求集成多个数据库，重新组织数据结构，统一规范编码，使其有效地完成各种决策分析。

从数据库到数据仓库的演变，体现了以下几点：（1）数据库用于事务处理，数据仓库用于决策分析。事务处理功能单一，数据库完成事务处理的增加、删除、修改、查询等操作。

决策分析要求数据较多。

数据仓库需要存储更多的数据，它不需要修改数据，它主要从大量数据中提取综合信息以及利用历史数据的规律得到预测信息。

（2）数据库保持事务处理的当前状态，数据仓库既保存过去的的数据又保存当前的数据。

数据库中的数据随业务的变化一直在更新，总保存当前的数据，如学生数据库、财务数据库等。

数据仓库中的数据不随时间变化而变化，但它保留大量不同时间的数据，即保留历史数据和当前数据。

（3）数据仓库的数据是大量数据库的集成。

数据仓库的数据不是数据库的简单集成，而是按决策主题，将大量数据库中的数据进行重新组织，统一编码进行集成。

如银行数据仓库数据是由储蓄数据库、信用卡数据库、贷款数据库等多个数据库按“用户”主题进行重新组织、编码和集成而建立的。

可见，数据仓库的数据量比数据库的数据量大得多。

（4）对数据库的操作比较明确，操作数据量少。

对数据仓库操作不明确，操作数据量大。

一般对数据库的操作都是事先知道的事务处理工作，每次操作（增加、删除、修改、查询）涉及的数据量也小，如一个或几个记录数据。

对数据仓库的操作都是根据当时决策需求临时决定而进行的。

如比较两个地区某个商品销售的情况。

该操作所涉及的数据量很大，不是几个记录数据，而是两个地区多个商店的某商品的所有销售记录。

## <<数据仓库与数据挖掘教程>>

### 编辑推荐

《数据仓库与数据挖掘教程(第2版)》特色：1.从数据库到数据仓库、从数据字典到元数据来说明数据仓库的兴起。

2.从联机事务处理(OLTP)到联机分析处理(OLAP)、多维数据的显示来说明多维数据的兴起。

3.通过实例来说明多维数据分析的应用与数据立方体新技术。

4.突出数据仓库的决策支持和基于数据仓库的决策支持系统。

5.从数据挖掘的理论基础来说明数据挖掘的方法：决策树方法的信息论原理、粗糙集方法和关联规则挖掘方法的集合论原理。

6.神经网络的介绍从简单到复杂；最早的感知机到反向传播网络BP，再到径向基函数网络。

通过神经网络几何意义的介绍，说明神经网络的分类效果。

7.变换规则的知识挖掘是在数据挖掘的基础上，增加规则前提的变换和结论的变换，获取变换规则知识。

变换规则知识是适应变化环境的新知识，是一般规则知识的发展。

8.软件进化规律的知识挖掘，系统地总结了软件进化过程和进化的规律，帮助读者提高软件开发能力

。

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>