

<<数据挖掘与数学建模>>

图书基本信息

书名：<<数据挖掘与数学建模>>

13位ISBN编号：9787118066715

10位ISBN编号：7118066710

出版时间：2010-2

出版时间：国防工业出版社

作者：廖芹，赫志峰，陈志宏 编著

页数：352

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>

## 前言

民众多好饮酒，中外概莫能外。

酒馆和酿酒坊伴随饮酒客而起，人类对酒的喜爱造就了酒文化和一个庞大的产业。

好酒能卖好价钱，能使文人诗兴大发，催生佳作，还能解人间百难。

于是，酿天下名酒自然成为不少人的毕生追求。

怎样才能酿出好酒呢？

国人的看法不尽相同。

崇信洋酒的人主张引进国外的生产工艺，学习洋人的生产和经营理念，而喜欢国酒的人则主张走自己的路，但不排除借鉴国外先进的科学技术和管理经验。

这样的争论或许永远不会终结，但外国人重视科学酿酒，这一点是值得我们学习和借鉴的。

计算机科学教育，如同酿酒工业的生产一样，科学办学迄今还只是部分学者的一种理想。

与国内一样，国外的计算机科学教育并没有像他们的科学酿酒业一样，实现科学办学。

也许科学办学要远比科学酿酒困难得多。

譬如，怎么实现科学办学？

甚至怎么推出一套科学的系列教材都是一篇大文章。

这套教材的创作始于教育部面向21世纪教育与教学改革13-22项目的研究。

2000年，在13-22项目研究工作即将完成之际，一些学者开始认识到面对计算机科学与技术的高速发展，我们亟需一套体现科学办学思想、反映内涵发展要求、服务教育与教学改革、参与构建学科人才培养科学体系的系列教材。

强调系列教材是因为那时已经意识到计算机科学教育本质上是一项科学活动，但长期以来教师向学生传授科学技术知识的方式方法科学性不强。

由于高等教育几百年来一直沿袭经验方式而非科学方式办学，大学教学的方式方法仍然还停留在古代作坊式的阶段，只不过今天使用的教学技术手段先进而已。

在经验办学方式下，无论是研究型大学还是教学型大学，由于种种原因，教学活动的全过程存在着太多的漏洞和质量上的隐患。

科学办学是对高等教育界传统的一个挑战，尽管在认识上，人们不难理解，科学办学是经验办学的最高形式，而经验办学应该成为科学办学的有益补充。

13-22项目组积极探索，率先倡导科学办学理念，初步构建了一个体现科学办学思想，反映内涵发展要求的计算机科学一级学科人才培养科学理论体系，为学科专业教育探索新天地，走向科学办学和发展学科系列教材提供了一个认知基础。

长期以来，学术界一直在探索计算机专业教育的规律。

## <<数据挖掘与数学建模>>

### 内容概要

本书以国际数据挖掘标准流程（GRISP-DM）为依据，以企业管理面临的现实问题为应用案例，由浅入深介绍数据挖掘方法及其解决问题过程的数据理解、数据处理、数据分析、数学建模、模型结果评估等内容，并引入应用广泛的数据挖掘Clementine软件辅助问题案例的解决，使读者不仅可以集中地学习数据挖掘的主要理论方法，而且可以了解基于数据挖掘的数学建模过程，可以学习应用软件辅助解决问题的操作方法。

本书把理论、案例、建模、软件辅助结合一体统一叙述，简述理论，突出应用，详细分析，展示过程，既考虑高校学生的学习需要，分本科生与研究生学习层次，又考虑企业管理者的应用与实践需要。

本书可作为数据挖掘理论与技术的教学、实践、应用和提高了的教科书或参考书。

适合高等学校本科高年级学生、研究生以及学习数据挖掘、数学模型课程的学生使用，也适合相关的企业管理与决策支持技术人员使用。

## &lt;&lt;数据挖掘与数学建模&gt;&gt;

## 书籍目录

第1章 数据挖掘与数学建模关系概述 1.1 当前信息化发展的趋势与面对问题 1.2 数据挖掘发展及其应用 1.3 基于数据挖掘的数学建模 1.4 数据挖掘软件Clementine的基本操作概述 参考文献 第2章 统计分析 2.1 问题概述 2.2 回归分析及其应用 2.3 二项逻辑回归 2.4 主成分分析 2.5 因子分析 2.6 管理胜任力的案例分析和数学建模 参考文献 第3章 聚类分析 3.1 问题概述 3.2 聚类分析概述 3.3 基于距离的聚类相似度 3.4 系统聚类法 3.5 C-均值(C-Means)聚类算法 3.6 Clementine辅助K-Means聚类 3.7 模糊聚类 3.8 聚类有效性 3.9 医疗建设评价的案例分析与数学建模 参考文献 第4章 神经网络及其应用 4.1 问题概述 4.2 神经网络概述 4.3 神经网络的基本模型 4.4 误差逆传播神经网络模型 4.5 RBF神经网络 4.6 自组织特征映射(SOM)神经网络 4.7 神经网络的案例分析与数学建模 4.8 模糊神经网络 参考文献 第5章 决策树及其应用 5.1 问题概述 5.2 决策树概述 5.3 ID3算法 5.4 C4.5算法 5.5 CART算法 5.6 Clementine辅助决策树建立 5.7 决策树的评价标准 5.8 决策树的剪枝 5.9 决策树的优化 5.10 燃气管网安全预测案例分析与数学建模 5.11 模糊决策树 参考文献 第6章 关联分析 6.1 问题概述 6.2 关联分析概述 6.3 Apriori关联规则算法 6.4 Clementine辅助Apriori算法实现关联分析 6.5 基于Apriori算法的改进算法 6.6 基于分类搜索的关联规则算法 6.7 基于频繁树的关联规则算法 6.8 关联规则的推广 6.9 时序关联规则算法 6.10 Clementine辅助时序关联分析 6.11 多值属性关联规则算法 6.12 增量关联规则算法 6.13 基于关联规则的分类算法 6.14 关联分类算法的改进 6.15 模糊关联分类算法与应用 6.16 关联规则案例分析与数学建模 参考文献 第7章 遗传算法 7.1 问题概述 7.2 遗传算法概述 7.3 模式定理 7.4 改进的遗传算法 7.5 遗传算法案例分析与数学建模 参考文献 第8章 贝叶斯网络与知识推理 8.1 问题概述 8.2 贝叶斯网络原理概述 8.3 主要贝叶斯网络模型 8.4 管理胜任力的贝叶斯网络知识推理模型建立 8.5 Clementine辅助贝叶斯网络的建立 8.6 地下燃气管网安全风险的贝叶斯网络知识推理与诊断 8.7 模糊贝叶斯网络 8.8 电子商务发展水平的模糊贝叶斯网络知识推理模型 8.9 数据挖掘软件WEKA辅助无约束贝叶斯网络的建立 参考文献

## 章节摘录

插图：如果检验结果达到期望的目标，表示模型建立完毕，如果检验发现有较大的偏差，即模型计算的目标值与实际样本目标值不一致，表示模型的结果不符合实际需要，此时要返回数据处理环节，检查提取样本特征和定义样本取值的合理性，当学习样本不足以代表检验样本信息或拟合模型偏离时，可能导致建立模型的失效性。

5) 结果评估建立的模型通过模型检验后，要对模型结果反映的数据关系的变化规律进行科学分析与评估，使与实际应用问题的物理意义与现实机理反映的特征规律相一致，如果模型结果反映的变化规律与实际的相不一致，表示尽管学习样本、检验样本与模型结果一致，但反映的变化规律仍与问题研究的目标有差异，此时应返回数据理解环节，分析可能存在但忽略的影响因素变量，通过补充，重新定义新的样本与模型结构，通过再学习使模型结果与实际问题的变化规律和物理意义相一致。

6) 结果应用当建立的数学模型不仅通过指定样本检验，而且通过应用问题的结果评估时，模型可以应用于问题领域，一般分为试应用和完善应用两个阶段。

由于基于数据挖掘的数学建模，是从数据切入，数据获取的部分性和定义样本的非全局代表性，可能使模型结果的应用也带有局部而非全局的特征，需要不断地扩大数据收集范围，扩大模型吸取学习样本的特征信息，才能扩展模型应用的全局性质，因此，试应用是不断补充样本特征信息使模型自适应学习修正模型参数的过程；完善应用是不断比较模型结果与专家经验，使模型结果与实际的期望结果更接近的完善过程。

1.4 数据挖掘软件Clementine的基本操作概述目前，各领域应用的数据挖掘软件辅助决策有许多，较常用的一个软件是由SPSS生产商推出的数据挖掘软件Clementine。

该软件按照国际定义的数据挖掘标准流程设计，包含了基本的数据挖掘方法，具有较强的数据与模型集成功能，使用方便直观，是一种较好的数据挖掘教学与应用分析工具。

1.4.1 Clementine数据挖掘的基本思想Clementine提供了大量的人工智能、统计分析的模型（神经网络，关联分析，聚类分析、因子分析等），并用基于图形化的界面为认识、了解、熟悉这个软件提供了方便。

除了这些，Clementine还拥有优良的数据挖掘设计思想，正是因为有了这个思想，每一步的工作也变得很清晰，如图1.1所示。

如图1.1所示，CRISP-DM Model包含了6个步骤，并用箭头指示了步骤间的执行顺序。

这些顺序并不严格，用户可以根据实际的需要反向执行某个步骤，也可以跳过某些步骤不予执行。

<<数据挖掘与数学建模>>

编辑推荐

《数据挖掘与教学建模》：现代计算机科学与技术教材系列。

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>