

<<现代统计研究丛著>>

图书基本信息

书名：<<现代统计研究丛著>>

13位ISBN编号：9787503756399

10位ISBN编号：750375639X

出版时间：2009-7

出版时间：吴喜之、马景义、吕晓玲、闫洁 中国统计出版社 (2009-07出版)

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>

## <<现代统计研究丛著>>

### 内容概要

《现代统计研究丛著:数据挖掘前沿问题》所涉及的是代表统计领域发展最快的部分,也是对传统统计提出最重大挑战的部分,这就是在数据挖掘实践中发展的新方法。而《现代统计研究丛著:数据挖掘前沿问题》所着重说明的前沿问题是:数据挖掘方法中最普遍应用的分类和回归中成为近年来发展热点的组合方法。

## 书籍目录

第1章 数据挖掘概论 1.1 引言 1.2 统计学家和计算机学家从不同角度看数据挖掘 1.3 数据源 1.4 数据挖掘的应用 第2章 传统统计面临的挑战 2.1 统计的黑匣子特性 2.2 统计从数学继承了什么 2.3 传统的数据建模在应用中所遇到的问题 2.4 算法建模 2.5 回到统计的最初宗旨 第3章 常用算法建模概述 3.1 引言 3.2 关联规则分析 3.3 最近邻方法 3.4 人工神经网络 3.5 支持向量机 3.6 VC维数和误差界限 第4章 决策树 4.1 引言 4.2 决策树的构建 4.3 不纯度 4.4 ID3和C4.5算法 4.5 CART算法 4.6 CHAID方法 第5章 模型评价 5.1 引言 5.2 贝叶斯规则 5.3 模型评价——再论CART 5.4 推广误差和期望推广误差 5.5 推广误差和期望推广误差的估计 第6章 Bagging预测方法 6.1 Bagging方法简介 6.2 分类问题的Bagging算法 6.3 回归问题的Bagging算法 6.4 Out—of—Bag (OOB) 估计 6.5 讨论 第7章 Boosting预测方法 7.1 AdaBoost算法 7.2 自适应重新抽样 7.3 AdaBoost算法的性质 7.4 可加模型：从统计的角度看AdaBoost 7.5 梯度下降提升算法 7.6 分类问题的不同损失函数及LogitBoost分类算法 7.7 回归问题的不同损失函数及L2—Boosting回归方法 7.8 讨论 第8章 随机森林 8.1 子模型 $h(x; m)$  8.2 随机森林用于分类的案例 8.3 分类问题中随机森林算法预测精度 8.4 随机森林算法用于回归问题 8.5 随机森林中的OOB估计 8.6 再析随机森林算法 8.7 自适应随机森林算法 参考文献

章节摘录

版权页：插图：由于很多数理统计课程基本上由数学老师教授，完全按照纯粹数学的模式设计，所以，对于背后的基于数据的统计思想介绍得不很充分，也不强调这些充满假定的数学模型都是对现实世界的不同程度的简化。

几乎没有人告诉学生，所有统计教科书中对数据（或其总体）的数学假定都是无法用数据验证的；大多数教科书仅仅指出这些模型在什么假定下可用，而很少指出违背这些假定的后果；统计教科书往往在给出统计方法结论的同时，不指出根据这些结论所作出决策的风险，也很少强调统计学家不能替代实际领域专家做决策的原则。

数学化的统计教科书极少提到统计应用中一系列决策的主观性和任意性。

2.3 传统的数据建模在应用中所遇到的问题 首先，无论是统计学家还是其他领域的研究人员，对他们的研究对象所选择的模型，无论是现成的，或者是他们要基于现成模型修正的，或者是他们针对这个课题所新建的，都仅仅是对现实世界的某种近似。

而这些用数学语言所描述的模型存在的一个必要条件是它们必须能够被人们解出来。

这些解可以是近似的，或者是精确的。

无论得到什么样的结论，都由于模型的近似性而必然是近似的。

而这些结果到底和现实世界有多么近似，则是不可能完全说清楚的。

## <<现代统计研究丛著>>

### 编辑推荐

《现代统计研究丛著:数据挖掘前沿问题》是国家社科基金重大项目，是一本现代统计研究丛著。

《现代统计研究丛著:数据挖掘前沿问题》主要研究的是关于数据挖掘前沿问题。

《现代统计研究丛著:数据挖掘前沿问题》讲述了一些关于数据挖掘的前沿统计方法，帮助读者学习数据挖掘，解决数据挖掘出现的问题。

<<现代统计研究丛著>>

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>