

大数据时代的统计与分析

第1章 大数据分析概述

1.1 大数据概述

1.1.1 什么是大数据

1.1.2 数据、信息与认知

1.1.3 数据管理与数据库

1.1.4 数据仓库

1.1.5 数据挖掘的内涵和基本特征

1.2 数据挖掘的产生与功能

1.2.1 数据挖掘的历史

1.2.2 数据挖掘的功能

1.3 数据挖掘与相关领域之间的关系

1.3.1 数据挖掘与机器学习

1.3.2 数据挖掘与数据仓库

1.3.3 数据挖掘与统计学

1.3.4 数据挖掘与智能决策

1.3.5 数据挖掘与云计算

1.4 大数据研究方法

第2章 数据挖掘流程

2.1 数据挖掘流程概述

2.1.1 问题识别

2.1.2 数据理解

2.1.3 数据准备

2.1.4 建立模型

2.1.5 模型评价

2.1.6 部署应用

2.2 离群点发现

2.2.1 基于统计的离群点检测

2.2.2 基于距离的离群点检测

2.2.3 局部离群点算法

2.3 不平衡数据级联算法

第3章 大数据的统计分析

3.1 相关分析

3.2 回归分析

3.3 决策树分析

3.3.1 决策树的基本概念

3.3.2 分类回归树

3.3.3 决策树的剪枝

3.4 假设检验

3.5 聚类分析

3.6 预测分析

3.7 单因子方差分析

第4章 大数据的统计分析案例

小概率的力量：航空安全与彩民信心

夜空中的灾难

4次灾难惊人的巧合

27000年才有一次的中奖机会

白点黑点，换个角度看数据

在整个背景下评价数据

精心选择的数字更丰富